Teaming up to keep tabs on a massive Amazon dam p. 18

Young scientists' favorite failures p. 24

Thousand Talents program nurtures Chinese scientists p. 62

Science

$15
6 JANUARY 2023
science.org
AAAS

MELTING AWAY

Glacial mass loss scales linearly with air temperature increase pp. 29 & 78

TRILLIONS OF MICROBES ONE ESSAY

The NOSTER *Science* Microbiome Prize is an international prize that rewards innovative research by investigators who have completed their terminal degree in the last 10 years and are working on the functional attributes of microbiota. The research can include any organism with the potential to contribute to our understanding of human or veterinary health and disease, or to guide therapeutic interventions. The winner and finalists will be chosen by a committee of independent scientists chaired by a senior editor of *Science*. The top prize includes a complimentary membership to AAAS, an online subscription to *Science*, and USD 25,000. Submit your research essay today.

Jennifer Hill, Ph.D.
2022 Winner

NOSTER | Science
MICROBIOME
PRIZE

Apply by 24 January 2023 at www.science.org/noster

Sponsored by Noster, Inc

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

CONTENTS

6 JANUARY 2023 • VOLUME 379
ISSUE 6627

18
Cleysson Juruna helps keep watch on how a giant hydroelectric project is affecting Brazil's Xingu River.

NEWS

IN BRIEF

8 Areas to watch: What to expect in research and policy

IN DEPTH

11 China is flying blind as the pandemic rages
Official death tolls are impossibly low, and some worry new variants may escape detection *By D. Normile*

12 Did ancient tentacled microbes kick-start complex life?
New studies suggest early Asgards evolved into eukaryotes *By E. Pennisi*

13 Once banned, spending earmarks see resurgence
Lawmakers can't resist steering cash to universities and research projects back home *By J. Mervis*

15 Ancient points suggest Asian roots for early American tools
Finds may support coastal route hypothesis for first settlers *By M. Price*

16 Virus hunters test new surveillance tools
Ropes, drones, insects, and dust cloths could make monitoring faster, safer, and cheaper *By J. Cohen*

FEATURES

18 A river's pulse
In the Brazilian Amazon, Indigenous people and researchers join hands to monitor the impacts of a controversial dam
By S. Moutinho and D. Galdieri
PODCAST

INSIGHTS

LETTERS

24 NextGen Voices: The fruits of failure

PERSPECTIVES

26 Wastewater surveillance for public health
Wastewater contains information on pathogen spread, evolution, and outbreak risk *By J. I. Levy* et al.

PHOTOS: (TOP TO BOTTOM) DADO GALDIERI/HILAEA MEDIA; SILVIO AVILA/AFP/GETTY IMAGES

28 Obesity confers macrophage memory
Epigenetic programming of myeloid cells in obesity contributes to macular degeneration
By K. D. Mangum and K. A. Gallagher
RESEARCH ARTICLE p. 45

29 Acting now will reduce glacier loss
Many of the world's glaciers will disappear, but quick action will make a difference
By G. Aðalgeirsdóttir and T. D. James
RESEARCH ARTICLE p. 78

30 Toward ultrathin optics
Function determines the minimum thickness of an optical system *By F. Monticone*
RESEARCH ARTICLE p. 41

POLICY FORUM

32 Soil biodiversity needs policy without borders
Soil health laws should account for global soil connections *By W. H. van der Putten* et al.

BOOKS *ET AL.*

35 The right to strive in a changing world
A philosopher prioritizes agency and Anthropocene concerns in a theory of animal justice *By D. Jamieson*

36 Materialism meets transcendence
A new documentary series follows a physicist's pursuit of the profound
By R. S. Krauss

RESEARCH

IN BRIEF

37 From *Science* and other journals

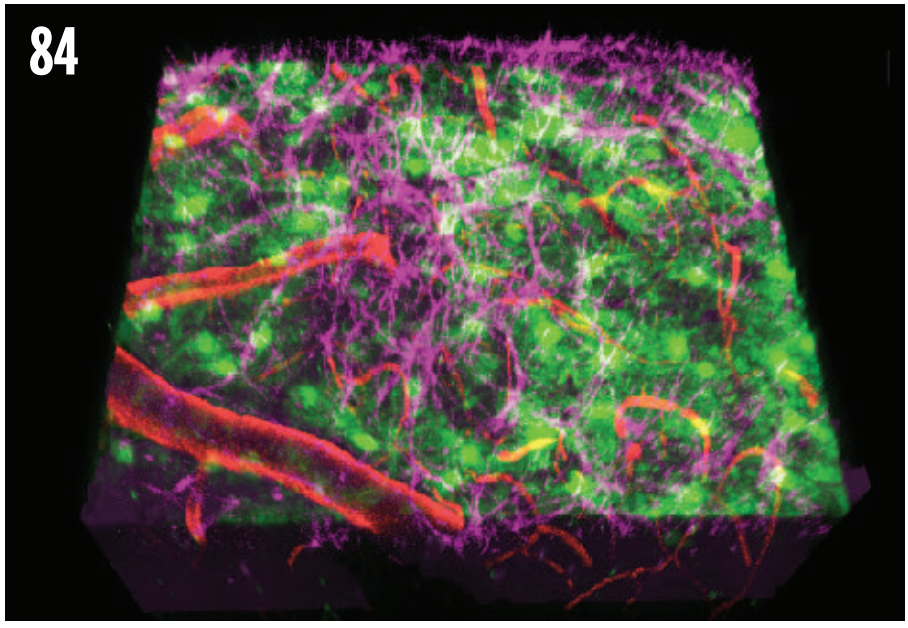
REVIEW

40 Optics
Lithium niobate photonics: Unlocking the electromagnetic spectrum *A. Boes* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABJ4396

RESEARCH ARTICLES

41 Optics
Why optics needs thickness *D. A. B. Miller*
PERSPECTIVE p. 30

45 Immunology
Past history of obesity triggers persistent epigenetic changes in innate immunity and exacerbates neuroinflammation *M. Hata* et al.
PERSPECTIVE p. 28



In vivo imaging reveals a fourth meningeal layer in the brain.

62 Expatriate scholars
Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists?
D. Shi et al.

Development
66 Immotile cilia mechanically sense the direction of fluid flow for left-right determination *T. A. Katoh* et al.

71 Cilia function as calcium-mediated mechanosensors that instruct left-right asymmetry *L. Djenoune* et al.

78 Glaciers
Global glacier change in the 21st century: Every increase in temperature matters
D. R. Rounce et al.
PERSPECTIVE p. 29

84 Brain anatomy
A mesothelium divides the subarachnoid space into functional compartments
K. Møllgård et al.

89 Metallurgy
Machine learning–aided real-time detection of keyhole pore generation in laser powder bed fusion *Z. Ren* et al.

94 Plant science
Maize resistance to witchweed through changes in strigolactone biosynthesis
C. Li et al.

100 Mpox
Structure of monkeypox virus DNA polymerase holoenzyme *Q. Peng* et al.

DEPARTMENTS

7 Editorial
Energy in 2023
By J. S. Yeston et al.

110 Working Life
Tired of persevering
By K. Hardeman

ON THE COVER

Meltwater from Nepal's Khumbu Glacier drains through an ice cave. By the end of the 21st century, 49 to 83% of the world's glaciers are projected to disappear, depending on future greenhouse gas emissions. Given that mass loss is directly related to temperature changes, every increase in temperature matters. As glaciers respond to climate change, societies must adapt to changes in water availability, natural hazards, and sea level. See pages 29 and 78. *Photo: Jason Gulley/American Reportage*

Science Staff 6
New Products 106
Science Careers 107

CREDIT: K. MØLLGÅRD *ET AL.*

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.

Copyright © 2023 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS.

Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2411; Foreign postage extra: Air assist delivery: $107. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science AAAS

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Valda Vinson

Editor, Research Jake S. Yeston **Editor, Insights** Lisa D. Chong **Managing Editor** Lauren Kmec

DEPUTY EDITORS Gemma Alderton (UK),Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORS** Caroline Ash (UK), Michael A. Funk, Brent Grocholski, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Mattia Maroso, Yevgeniya Nusinovich, Ian S. Osborne (UK), L. Bryan Ray, Seth Thomas Scanlon (UK), H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible **ASSOCIATE EDITORS** Bianca Lopez, Madeleine Seale (UK), Corinne Simonti, Yury V. Suleymanov, Ekeoma Uzogara **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Chris Filiatreau, Harry Jach **SR. CONTENT PRODUCTION EDITOR** Amelia Beyna **CONTENT PRODUCTION EDITORS** Robert French, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SR. EDITORIAL MANAGERS** Carolyn Kyle, Beverly Shields **SR. PROGRAM ASSOCIATE** Maryrose Madrid **EDITORIAL ASSOCIATE** Joi S. Granger **SR. EDITORIAL COORDINATORS** Aneera Dobbins, Jeffrey Hearn, Lisa Johnson, Jerry Richardson, Hilary Stewart (UK), Alice Whaley (UK), Anita Wynn **EDITORIAL COORDINATORS** Maura Byrne, Alexander Kief, Ronmel Navas, Isabel Schnaidt, Qiyam Stewart, Brian White **RESEARCH & DATA ANALYST** Jessica L. Slater **ADMINISTRATIVE COORDINATOR** Karalee P. Rogers **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Shraddha Chakradhar, Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Michael Price, Kelly Servick **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Zack Savitsky, Viviana Flores, Katherine Irving **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Tokyo), Elisabeth Pain (Careers), Charles Piller, Gabriel Popkin, Michael Price, Joshua Sokol, Richard Stone, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Morgan Everett, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Chrystal Smith **GRAPHICS MANAGING EDITOR** Chris Bickel **PHOTOGRAPHY MANAGING EDITOR** Emily Petersen **MULTIMEDIA MANAGING PRODUCER** Kevin McLean **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Marcy Atarod **DESIGNER** Christina Aycock **SENIOR SCIENTIFIC ILLUSTRATOR** Valerie Altounian **SCIENTIFIC ILLUSTRATORS** Austin Fisher, Kellie Holoski, Ashley Mastin **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Charles Borst **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark **FINANCIAL ANALYST** Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **SENIOR PRODUCTION MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **CONTENT OPERATIONS MANAGER** Rebecca Doshi **SENIOR CONTENT & PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wowk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Aaron Helmbrecht, Ashley Hylton, Mike Romano, Lorena Chirinos Rodriguez, Jenna Voris **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst **PROJECT MANAGER** Melissa Collins

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATE** Caroline Breul, Anne Mason **SPJ ASSOCIATE MANAGER** Samantha Bruno Fuller **SPJ ASSOCIATE** Casey Buchta

MARKETING MANAGER Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Aminta Reyes

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Rie Rambelli (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

MAIN HEADQUARTERS
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
science.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS
science.org/help/
reprints-and-permissions

MEDIA CONTACTS
scipak@aaas.org

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.science.org/
products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.science.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/benefits

AAAS BOARD OF DIRECTORS
CHAIR Susan G. Amara
PRESIDENT Gilda A. Barabino
PRESIDENT-ELECT Keith Yamamoto
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Ann Bostrom
Janine Austin Clayton
Kaye Husbands Fealing
Maria M. Klawe
Jane Maienschein
Robert B. Millard
Babak Parviz
William D. Provine
Juan S. Ramírez Lugo
Susan M. Rosenberg

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Erin Adams, *U. of Chicago*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Fdn.*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
James Analytis, *UC Berkeley*
Paola Arlotta, *Harvard U.*
Delia Baldassarri, *NYU*
Nenad Ban, *ETH Zürich*
Christopher Barratt, *U. of Dundee*
Nandita Basu, *U. of Waterloo*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *UT Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Kiros T. Berhane, *Columbia U.*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med.*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Malcolm Brenner, *Baylor Coll. of Med.*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
Johannes Buchner, *TUM*
György Buzsáki, *NYU School of Med.*
Mariana Byndloss, *Vanderbilt U. Med. Ctr.*
Annmarie Carlton, *UC Irvine*
Simon Cauchemez, *Inst. Pasteur*
Ling-Ling Chen, *SIBCB, CAS*
Wendy Cho, *UIUC*
Ib Chorkendorff, *Denmark TU*
Chunaram Choudhary, *Københavns U.*
Karlene Cimprich, *Stanford U.*
Laura Colgin, *UT Austin*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Virginia Cornish, *Columbia U.*
Carolyn Coyne, *Duke U.*
Roberta Croce, *VU Amsterdam*
Christina Curtis, *Stanford U.*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *UNC*
Nicolas Dauphas, *U. of Chicago*
Frans de Waal, *Emory U.*
Claude Desplan, *NYU*
Sandra Díaz, *U. Nacional de Córdoba*
Samuel Díaz-Muñoz, *UC Davis*
Ulrike Diebold, *TU Wien*
Stefanie Dimmeler, *Goethe-U. Frankfurt*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Ruth Drdla-Schutting, *Med. U. Vienna*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Nader Engheta, *UPenn*
Karen Ersche, *U. of Cambridge*
Beate Escher, *UFZ & U. of Tübingen*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Natascha Förster Schreiber, *MPI Extraterrestrial Phys.*
Peter Fratzl, *MPI Potsdam*
Elaine Fuchs, *Rockefeller U.*
Daniel Geschwind, *UCLA*
Ramon Gonzalez, *U. of South Florida*
Gillian Griffiths, *U. of Cambridge*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Taekjip Ha, *Johns Hopkins U.*
Daniel Haber, *Mass. General Hos.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *UCL*
Carl-Philipp Heisenberg, *IST Austria*
Janet G. Hering, *Eawag*
Christoph Hess, *U. of Basel & U. of Cambridge*
Heather Hickman, *NIAID, NIH*
Hans Hilgenkamp, *U. of Twente*
Janneke Hille Ris Lambers, *ETH Zürich*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Gwyneth Ingram, *ENS Lyon*
Darrell Irvine, *MIT*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS & U. of Arizona*
Erich Jarvis, *Rockefeller U.*
Peter Jonas, *IST Austria*
Johanna Joyce, *U. de Lausanne*
Matt Kaeberlein, *U. of Wash.*
William Kaelin Jr., *Dana-Farber*
Daniel Kammen, *UC Berkeley*
Kisuk Kang, *Seoul Nat. U.*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med.*
Nancy Knowlton, *Smithsonian*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
LaShanda Korley, *U. of Delaware*
Paul Kubes, *U. of Calgary*
Chris Kuzawa, *Northwestern U.*
Laura Lackner, *Northwestern U.*
Gabriel Lander, *Scripps Res.* **(S)**
Mitchell A. Lazar, *UPenn*
Hedwig Lee, *Duke U.*
Ryan Lively, *Georgia Tech*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *WUSTL*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Tech*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Zeynep Madak-Erdogan, *UIUC*
Anne Magurran, *U. of St. Andrews*
Ari Pekka Mähönen, *U. of Helsinki*
Asifa Majid, *U. of Oxford*
Oscar Marín, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth*
Rodrigo Medellín, *U. Nacional Autónoma de México*
C. Jessica Metcalf, *Princeton U.*
Tom Misteli, *NCI, NIH*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Thi Hoang Duong Nguyen, *MRC LMB*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci. & Tech. Fund*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *MPI Cologne*
Giovanni Parmigiani, *Dana-Farber* **(S)**
Daniel Pauly, *U. of British Columbia*
Ana Pêgo, *U. do Porto*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Inst. Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Lei Stanley Qi, *Stanford U.*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial Coll. London*
Amy Rosenzweig, *Northwestern U.*
John Rubenstein, *SickKids*
Mike Ryan, *UT Austin*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Erica Ollmann Saphire, *La Jolla Inst.*
Joachim Saur, *U. zu Köln*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UCSF*
Robert Siliciano, *JHU School of Med.*
Lucia Silvotti, *UCL*
Emma Slack, *ETH Zürich & U. of Oxford*
Richard Smith, *UNC* **(S)**
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Northwestern U.*
Ira Tabas, *Columbia U.*
Eriko Takano, *U. of Manchester*
A. Alec Talin, *Sandia Natl. Labs*
Patrick Tan, *Duke-NUS Med. School*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Maria-Elena Torres Padilla, *Helmholtz Zentrum München*
Kimani Toussaint, *Brown U.*
Barbara Treutlein, *ETH Zürich*
Jason Tylianakis, *U. of Canterbury*
Wim van der Putten, *Netherlands Inst. of Ecology*
Matthew Vander Heiden, *MIT*
Ivo Vankelecom, *KU Leuven*
Judith Varner, *UC San Diego*
Henrique Veiga-Fernandes, *Champalimaud Fdn.*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Julia Von Blume, *Yale School of Med.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Hao Wu, *Harvard U.*
Li Wu, *Tsinghua U.*
Wei Xie, *Tsinghua U.*
Benjamin Youngblood, *St. Jude*
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *UPenn School of Med.*
Lidong Zhao, *Beihang U.*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

EDITORIAL

Energy in 2023

Jake S. Yeston
Editor of Research, *Science*. jyeston@aaas.org

Valda Vinson
Executive Editor, *Science*. vvinson@aaas.org

H. Holden Thorp
Editor-in-Chief, *Science* journals. hthorp@aaas.org; @hholdenthorp

The new year 2023 arrives with promising developments in fusion research. In December, scientists at the US National Ignition Facility (NIF) focused 2.05 megajoules of laser light onto a capsule of fusion fuel and produced 3.15 megajoules of energy. This was the first laser-driven fusion demonstration in which the reaction produced more energy than the laser light used to start it—a goal set for NIF at its founding. Massachusetts Institute of Technology physicist Anne White, who was not involved in the work, described it to *Science* as "a major breakthrough." After a year of ups and downs in science policy and research, this was the kind of exciting event needed to propel scientists into the promises of 2023.

As always, with these kinds of advances, scientists and the federal government must be careful not to exaggerate the practical implications. The achievement was announced at a high-profile press conference hosted by the US Secretary of Energy. But as *Science*'s Daniel Clery wrote, "fusion power stations are still a distant dream." The NIF result is a glimmer of promise, but many technical challenges still lie ahead. Scientific results announced by press conference prior to the publication of a peer-reviewed paper do not always hold up well, so we will have to wait to see the data, hopefully this year.

"We are making a number of important changes to our Information for Authors and Editorial Policies."

Similarly, 2023 brings new energy to *Science*. We are making a number of important changes to our Information for Authors and Editorial Policies. These include updating the types of research papers that we publish and adding a new category for retractions.

Over the history of the magazine, research papers have gotten longer and longer with more data and figures. These developments are good for open science, as the experimental methods and results must be reported in sufficient detail to allow reproducibility. As a result, *Science* is ending the category of Reports, which were shorter than Research Articles. Now there will be one category of research paper—the Research Article—which will be around five pages in the printed journal. This will typically include 2000 to 3000 words of text, three to five figures, and 50 references. The majority of *Science*'s papers already fit into this category. This format still requires the focused writing that enhances the clarity of the published papers, but without compromising transparent reporting on the work. Supplementary Materials will continue to be published online, with the accompanying data, curated at repositories such as Zenodo and Dryad, expected to grow. We will continue to allow some longer Research Articles to appear online and with a one-page summary in the print magazine.

Up until now, two criteria for retractions have been explained in the Editorial Policies. One applies to instances when misconduct is clearly demonstrated, and the other concerns errors that undermine the core conclusions of a paper. All other errors were often accommodated by *Science*'s publication of an Erratum or a Corrections notice. We are now adding a third criterion for when a paper has received enough corrections or contains enough errors to cause the editors to lose confidence in it. Although *Science* has always had the latitude to retract a paper for reasons deemed sufficient to the editors, we are now stating this additional criterion explicitly.

The analysis of published images in research papers has become more prominent and cataloged on sites like PubPeer, and image sleuths like Elisabeth Bik have brought more energy to the careful maintenance of the scientific record. Scientists are human, and mistakes—whether intentional or inadvertent—are inevitable by both authors and journals. Fortunately, science is a self-correcting process that ultimately leads to the right answer. Unfortunately, when we contact authors about concerns that have been raised about their *Science* papers, we are often met with defensiveness and denial. That needs to change. A quick correction or retraction builds confidence that we all are committed to an accurate scientific record. The community needs to work to reduce the stigma of correcting the record so that—in this time of extraordinary scrutiny and waning confidence—we can show the world that self-correction provides a basis for trust. We hope that this new criterion for retractions will contribute to enhancing that trust.

Here's to an energetic 2023—for fusion and more.

–Jake S. Yeston, Valda Vinson, H. Holden Thorp

10.1126/science.adg4111

NEWS

IN BRIEF

Edited by **Jeffrey Brainard**

A crowd boards a bus in Germany, where some states have eased masking rules for passengers on local public transit.

PHOTO: JULIAN RETTIG/PICTURE ALLIANCE/DPA/AP IMAGES

AREAS TO WATCH: WHAT TO EXPECT IN RESEARCH AND POLICY

Science strives to keep up with an evolving coronavirus

As the COVID-19 pandemic enters its fourth year as a global health emergency, researchers will continue pushing to help make the disease manageable and ordinary. They will track hundreds of subvariants of Omicron, the highly transmissible but seemingly less lethal strain of SARS-CoV-2 that dominated in 2022. Virologists will watch the virus' evolution this year to see whether it has finally slowed or a more dangerous variant pops up, evading much of the immunity that humanity has built up to previous ones. Vaccine researchers hope to develop new shots that provide broad protection against a variety of coronaviruses. Another priority is to introduce nasal vaccines that prompt immune responses within the body's mucosal linings; compared with shots in the arm, these should elicit a stronger, quicker defense against initial infection. But public health specialists worry widespread vaccine hesitancy may persist, with long-lasting consequences for battles against both COVID-19 and other diseases. The search for effective treatments for COVID-19 will reboot because the evolution of the virus has made several existing antibody-based drugs ineffective. Randomized trials of potential drugs to treat Long Covid may yield their first results, benefiting millions of people suffering from fatigue and other debilitating symptoms.

In the rest of this section, *Science*'s News staff forecasts other areas of research and policy likely to make headlines in the coming year.

Nations mull climate-losses pact

CLIMATE POLICY | This year, diplomats from two dozen countries will debate the terms of an agreement requiring wealthy nations, responsible for most historical greenhouse gas emissions, to help pay for damages caused by climate change. The commitment was the only substantive new policy to emerge from the U.N. climate summit in Egypt in November 2022. A new fund would pay for economic losses and property damage linked to heat, flooding, and other effects of climate change. But signatories put off details, such as which countries should pay, which should benefit, and how the money should be spent. Those details could come into focus at this year's U.N. climate summit in the United Arab Emirates. Observers are skeptical about the deal's prospects, noting that wealthy nations have failed to honor past promises to provide such financial support.

Big funders to get new leaders

BIOMEDICINE | Two of the world's largest biomedical research sponsors—the U.S. National Institutes of Health (NIH) and the Wellcome Trust—will likely get new directors this year. Researchers have grown impatient waiting for the White House to nominate a successor to Francis Collins, who stepped down as director of the $47.5-billion-a-year NIH in December 2021. The agency's next leader, who will require Senate confirmation, will oversee NIH's efforts to boost diversity in the research workforce and likely face a grilling by Republican lawmakers over the agency's response to the COVID-19 pandemic. Also awaiting a permanent director is NIH's $6.6 billion National Institute of Allergy and Infectious Diseases, which Anthony Fauci led for 38 years before he stepped down last month. In addition, Wellcome, a nonprofit organization that provides more than £1 billion annually for research, will seek a replacement for Director Jeremy Farrar, who announced he will leave early this year after a decade in that role to become chief scientist of the World Health Organization.

Human gene editing nears OK

BIOMEDICINE | The new year could bring a milestone for medicine: the first approved medical treatment based on gene editing. People with sickle cell disease and beta-thalassemia carry defects in the gene for hemoglobin, the oxygen-carrying protein in blood. The biotech companies Vertex Pharmaceuticals and CRISPR Therapeutics have run clinical trials in which they remove a patient's blood stem cells, use the CRISPR gene-editing tool to switch on a healthy gene for fetal hemoglobin, which cells shut off after birth, and reinfuse the modified cells. The one-time treatment has ended most patients' severe pain episodes and the need for blood transfusions. The companies are seeking approval from U.S. and European regulators, and a decision on at least one side of the Atlantic could come by year's end. The next concern will be cost. Gene therapy, an older approach that treats genetic disorders by adding rather than modifying genes, has come with price tags from $850,000 to $3.5 million.

Animal genomes proliferate

GENOMICS | Scientists have sequenced and studied thousands of human and microbial genomes, but the complete deciphering of the DNA of other multicellular organisms has lagged. A new chapter will begin this year, when a surge of nonhuman genome sequences will be unveiled, the fruits of cheaper, more precise technologies. The Earth BioGenome Project, an umbrella effort that encompasses more than 50 sequencing efforts covering a broad range of organisms, expects to see 2000 sequences released this year. Many groups under that umbrella have focused on particular types of animals: For instance, the Wellcome Sanger Institute's Darwin Tree of Life Project has been studying insects and other invertebrates, and the Zoonomia Project has scrutinized mammals. Two other consortia will release sequences of more than 200 nonhuman primates, many of which have been designated threatened or endangered. Scientists expect the new data will boost comparative genomics studies and provide insights into evolution and life history.

A bid to end mpox

PUBLIC HEALTH | Health authorities will strive this year to eliminate human-to-human transmission of mpox (previously known as monkeypox), which in 2022 exploded across the globe for the first time. More than 80,000 people were sickened, prompting the World Health Organization to declare a Public Health Emergency of International Concern. As the year ended, the number of new cases had declined steeply, but hundreds of cases were still reported every week. Experts hope to better understand how much of the decline stems from immunity built up through infections and immunizations and how much can be explained by behavior change in the most affected group: gay men and their sexual networks. Studies underway could reveal how well the one vaccine, a repurposed smallpox shot, protects against the disease and compare different doses and ways to administer the shots. Whether the vaccine will be available in African countries that have long experienced mpox outbreaks will be a crucial test of global health policy.

China to launch space telescope

ASTRONOMY | China's space science efforts will continue to mature with the launch of three missions this year, including the Chinese Survey Space Telescope, also called Xuntian ("sky survey"). The 2-meter

PHOTO: MARCELO ISMAR SANTANA/UNIVERSITY OF BRAZIL

The Vieira's titi monkey (*Plecturocebus vieirai*) in Brazil is among many endangered species of primates whose genomes will be published, in one of several projects that have been sequencing the DNA of animal species.

telescope, China's first space-based optical probe, is not quite as big as the 2.4-meter Hubble Space Telescope, but it will have a field of view 350 times larger, benefiting astronomical surveys. It will study exoplanets, star formation, and galaxy evolution, as well as dark matter and dark energy, mysterious phenomena that control the expansion of the universe. The telescope will fly in the same orbit as China's Tiangong-3 space station and be able to dock for maintenance. Xuntian is scheduled to launch in December and start observations in 2024. The Chinese Academy of Sciences's National Astronomical Observatories has promised to share the data. China's other two missions are x-ray observatories—the Space Variable Objects Monitor, a joint project with France, and the Einstein Probe, in cooperation with the European Space Agency.

Anthropocene may get a marker

GEOLOGY | Researchers will soon announce their choice of a site to serve as the "golden spike" for the Anthropocene epoch, a controversial proposal to designate an official geological span of time marked by humanity's indelible effects on the planet. The Anthropocene Working Group, assembled by the International Commission on Stratigraphy, already picked the 1950s, an era of surging fossil fuel use, as the start of the epoch. And the group has reviewed 12 candidate sites around the world that contain lake muds, ice cores, or other features that clearly document the shift in emissions and could be used to formally define the new epoch. After the working group announces its choice, three more committees will have to approve it. Should the definition be voted down, it cannot be reconsidered for 10 years. Passage is far from guaranteed. Many geologists acknowledge the unprecedented changes wrought by human activity but question the need to change a system that describes millions of years of geologic time to mark shorter lived events driven by humanity.

Dengue vaccine nears debut

INFECTIOUS DISEASES | A new vaccine against dengue won approval in Europe last month and may soon become widely available in Indonesia, protecting far more people than a product currently on the market. Only people previously infected by the virus can safely take a dengue vaccine marketed since 2015, Sanofi Pasteur's Dengvaxia. Now, the pharmaceutical company Takeda has shown in multicountry studies of more than 28,000 people that its vaccine, Qdenga, can safely protect people never infected by dengue. The virus causes fever and other debilitating symptoms in an estimated 100 million people a year, and, in rare cases, can be fatal. Some scientists want to see more data about Qdenga's safety. Their wariness comes from experience with Dengvaxia: Children who had never been infected by the virus and received the vaccine had a higher risk of severe symptoms. This problem—which may be linked to an unusual phenomenon in which antibodies to one group of dengue viruses "enhance" the ability of a distinct second group to infect cells—has not occurred in Qdenga trials.

Researchers remove a sediment core at Crawford Lake in Canada, a site under consideration to mark humanity's geological footprint.

PHOTO: BRENNA BARTLEY/CONSERVATION HALTON

Amazon conservation reboots

LEADERSHIP | Brazil's new left-wing president, Luiz Inácio "Lula" da Silva, is expected this year to renew efforts to protect the Amazon and fight climate change—a reversal from the pro-development agenda of his far-right predecessor, Jair Bolsonaro, who presided over record deforestation. Lula, who was also president from 2003 to 2010, has promised to end illegal deforestation. He may start by revoking hundreds of laws and decrees issued by Bolsonaro that weakened policing of illegal logging and mining. Enforcement may be limited by other demands on Brazil's budget, however, such as fighting rising poverty. But other countries might help by reinstating subsidies to support conservation that were suspended during Bolsonaro's term. Norway and Germany both resumed such contributions days after Lula's election in October 2022.

South Africa tells *Homo*'s story

HUMAN ORIGINS | New analyses this year could lend support to the idea that key events in the evolution of our genus, *Homo*, happened in South Africa. Researchers working in Kromdraai Cave say they plan to publish descriptions of newly unearthed fossils of *Homo* that may date to earlier than 2 million years ago—soon after the date of the earliest *Homo* fossils, 2.7 million years ago in east Africa. Distinctive features of the Kromdraai fossils, including rarely found lower limbs, could bolster a controversial hypothesis that fossils discovered last year at nearby Drimolen quarry belonged to *H. erectus*, a direct human ancestor—which might indicate that the species first appeared in South Africa rather than in east Africa or Asia, as many have thought. Analyses of other new South African fossils, including forerunners of *Homo*, could also help untangle the histories and relationships of hominin species that lived in the area.

SCIENCE.ORG/NEWS
Read more news from *Science* online.

IN DEPTH

Tongren Hospital in Shanghai was overflowing with patients on 3 January. Even basic epidemiological data about China's COVID-19 wave are lacking.

PHOTO: HECTOR RETAMAL/AFP/GETTY IMAGES

COVID-19

China is flying blind as the pandemic rages

Official death tolls are impossibly low, and some worry new variants may escape detection

By **Dennis Normile**

Most scientists believe China's decision to end its zero-COVID policy was long overdue. But now they have a new worry: that the country is collecting and sharing far too little data about the rough transition to a new coexistence with the virus.

China abruptly dropped virtually all controls a month ago, after protests, a sagging economy, and the extreme transmissibility of the virus' latest variants made clinging to zero COVID untenable. Now, "SARS-CoV-2 has an open goal in front of it: a population with very low levels of standing immunity," says evolutionary biologist Edward Holmes of the University of Sydney. But how the epidemic is unfolding is a mystery because the country has practically stopped collecting basic epidemiological data.

Models that predicted a massive wave of infection and death if China ended zero COVID appear to have been correct. Press reports and social media posts have shown intensive care units stretched beyond capacity, with crowds of patients in wheelchairs and on gurneys in hallways. Doctors and nurses are reportedly working while sick. Crematoriums are overwhelmed. But China's official COVID-19 death toll is widely considered laughably low. And some scientists worry a genomic monitoring plan unveiled last month doesn't have the power to detect new SARS-CoV-2 variants arising as the virus works its way through one-fifth of the world population.

Earlier in the pandemic China's daily counts of COVID-19 cases and deaths, based partly on its exhaustive testing programs, were generally believed to be accurate. Now, they are anyone's guess. Patients with mild symptoms are not encouraged to get tested, let alone those who are asymptomatic. People testing positive at home are not asked to report their results.

The China Center for Disease Control and Prevention (China CDC) put the number of confirmed cases the last week of December 2022 at more than 35,000—a fraction of the official number in the United States. But leaked notes from an internal meeting suggest a very different reality: The agency was told that almost 250 million people in China—roughly 18% of the population—may have caught COVID-19 in the first 20 days of December. Some experts said the number is implausibly large, but Yanzhong Huang, a global health specialist at the Council on Foreign Relations, a U.S. think tank, says it's "not unreasonable," given credible reports that 80% of Beijing residents have been infected by now.

As to the death toll, China's reporting had long been inconsistent, Huang says, with some regions reporting all fatalities in which SARS-CoV-2 was a factor, as most countries do, and others excluding people who died from other conditions, such as heart attacks, even if they had COVID-19. In early December, China's government decided the narrower definition should be used nationwide.

Even then, the official count is astonishingly low: just eight deaths for the entire last week of December, which is "not matching media reports and what is being seen on social media," says Louise Blair, who tracks China's COVID-19 outbreak for Airfinity, a London-based health analytics firm that estimates about 9000 people were dying of COVID-19–related causes every day in late December. Also missing are data on case fatality rates, the average number of new infections stemming from each case, and hospital and intensive care admissions. "These are critical data" that would help health authorities get a handle on the surge and further the world's understanding of the pandemic, says Xi Chen, a public health scientist at the Yale School of Public Health.

A major worry is that the wave will breed a new and even more troublesome SARS-CoV-2 variant. "It's possible that something might be emerging, because there is such a big population in China," says George Gao, who in July 2022 stepped down as head of

China CDC but is now helping track circulating variants. But, he told *Science*, "There are no novel mutants—yet." At a 20 December press briefing, Xu Wenbo, head of the National Institute for Viral Disease Control and Prevention, explained that the BA.5.2 and BF.7 Omicron subvariants, which are now causing most infections globally, are also dominant in China. BQ.1 and XBB, which have recently been spreading in Europe and North America, have turned up in limited numbers in several provinces.

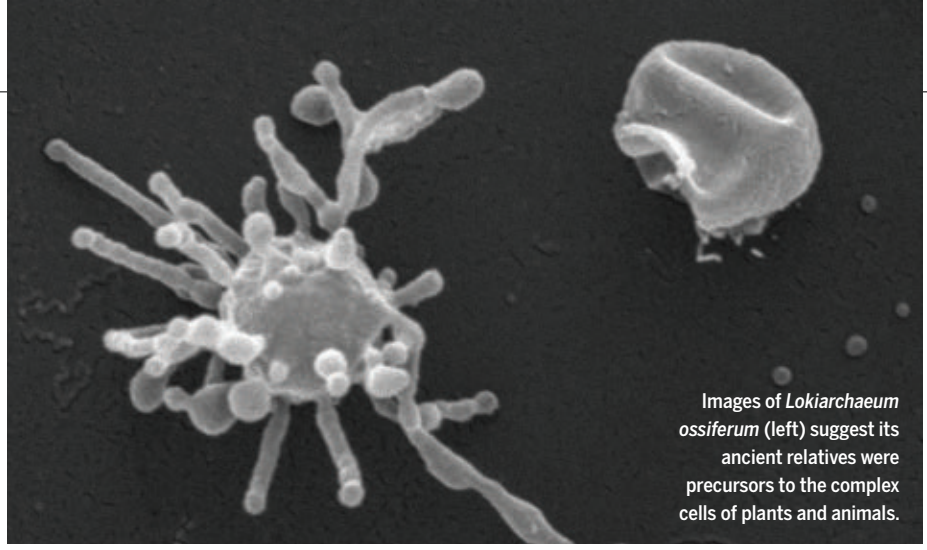
Experts are split on whether China is looking closely enough. Three designated sentinel hospitals in different cities in each of China's 31 provinces, municipalities, and regions are supposed to sequence and analyze samples from 15 outpatients, 10 severe cases, and all deaths every week. "I'm afraid [that] sample size is too small," Chen says. A stronger plan would consider province size and population density, instead of picking three cities in each, and adopt other sampling approaches, says Elizaveta Semenova, an epidemiologist at the University of Oxford. Semenova is a co-author of a study of how well 189 countries have detected new variants, published in November 2022. It concluded that effective surveillance requires sequencing about 0.5% of cases, with a turnaround time of less than 21 days. China's plan is unlikely to come close to that percentage.

But Gao and others say the program will pick up new variants in a timely way. If the surveillance plan gets up to speed, it will sequence 2000 to 3000 genomes per week, a level that "should be able to detect [new variants] and their transmission trends," says Leo Poon, a virologist at the University of Hong Kong who has helped track variants there. Poon notes that China is sharing data from the outbreak on GISAID, the world's largest database of SARS-CoV-2 sequences.

In addition to China CDC, research groups at more than 30 hospitals and universities are also tracking SARS-CoV-2 variants, says a Chinese epidemiologist who asked not to be identified. These groups "will report immediately if a dangerous variant emerges," the source says.

Still, the lack of reliable data is already undermining faith in China's handling of the outbreak. A dozen countries, including the United States and France, have announced they will require pre- or postflight tests on air travelers from China. Huang says that is unlikely to keep new variants out. The goal should be to convince the Chinese to be more forthcoming about what's happening on the ground—and for that, "quiet diplomacy may work better than travel restrictions," he says. ■

With reporting by Bian Huihui.



Images of *Lokiarchaeum ossiferum* (left) suggest its ancient relatives were precursors to the complex cells of plants and animals.

PHOTO: THIAGO RODRIGUES-OLIVEIRA/UNIVERSITY OF VIENNA

EVOLUTION

Did ancient tentacled microbes kick-start complex life?

New studies suggest early Asgards evolved into eukaryotes

By **Elizabeth Pennisi**

What may be the earliest ancestors of humans, plants, and all other complex life are coming out of the shadows. The unusual tentacled microbes called Asgards were initially known just from DNA extracted from marine sediments. Then, one was grown a few years ago in the lab—but in such low numbers that only its genome could be studied in detail. Now, recent studies are painting a clearer portrait of these organisms and strengthening the case that the complex cells making up all multicellular life—and some single-cell organisms—are descendants of ancient Asgard-like microbes.

Some recent work has shown that Asgards, which belong to a group of rudimentary life forms known as archaea, have genes once thought to exist only in more complex organisms. And in the most striking study, published last month in *Nature*, researchers cultured a second Asgard, found in mud from an estuary in Slovenia. They managed to grow enough for individual microbes to be imaged with electron microscopy, revealing complex internal structures suggestive of those in our own cells, such as the complex cytoskeleton made of the protein actin.

"This paper is beautiful," says Buzz Baum, an evolutionary cell biologist at the Medical Research Council's Laboratory of Molecular Biology in London. "The images are stunning."

Considered a third domain of life by most scientists, archaea are distinct from bacteria and eukaryotes, the evolutionary branch that encompasses everything from microbes such as yeast to humans. Still, archaea and bacteria have key similarities: Both generally lack core eukaryotic features such as mitochondria, cells' internal powerhouses, or DNA encased inside a nucleus. Many researchers think early eukaryotic cells arose some 2 billion years ago when an archaeon engulfed bacteria, which eventually became the mitochondria. But they have struggled to explain how other features of eukaryotes evolved, such as their many internal structures. "Until recently, life's journey towards complexity was a blur," says Masaru Nobu, a microbiologist at Japan's National Institute of Advanced Industrial Science and Technology.

The idea that Asgard-like archaea might be the ancestors of eukaryotes emerged in 2015. Thijs Ettema, an environmental microbiologist at Wageningen University, discovered eukaryotelike genes in archaeal genetic material from sediment samples collected by Christa Schleper, now a microbiologist at the University of Vienna, and her student, Steffen Jørgensen. By 2017, Ettema had found similar genes in several more groups of archaea, which he collectively called the Asgards, after the home of the Norse gods.

At the time, however, Ettema knew the Asgards only from roughly assembled genomes cobbled together from environmental DNA (eDNA)—the jumble of genetic material from many organisms in a soil or water sample. Skeptics argued he couldn't be sure the eukaryotelike genes really belonged to archaea. But in 2019, Nobu, with microbio-

logist Hiroyuki Imachi of the Japan Agency for Marine-Earth Science and Technology and colleagues, cultured the first Asgard microbe, isolated from ocean mud off Japan, and reported that its genome also had eukaryotic genes.

Additional evidence came earlier this year when Victoria Orphan, a geobiologist at the California Institute of Technology, and her colleagues isolated enough of two other Asgard species—this time from a hydrothermal vent in the Gulf of California—to sequence their complete genomes. Their genes also suggested a kinship with eukaryotes, Orphan and her colleagues reported on 13 January 2022 in *Nature Microbiology*.

Researchers including Ettema, Baum, and Mohan Balasubramanian, a cellular microbiologist at the University of Warwick, recently pinpointed a more specific link. They scanned the Asgard genomes sequenced so far for genes encoding the protein complexes that eukaryotic cells use to shape and stitch together membranes to form internal compartments. Known archaea lack those same compartments and were believed to have no more than two of the protein complexes. But Asgard genomes contain instructions for making four of them, the team reported on 13 June 2022 in *Nature Communications*, just one short of five complexes in eukaryotic cells.

In the lab, the Asgard proteins appeared to work similarly to the eukaryotic versions. To the scientists, that suggests the membrane-manipulating machinery predates the evolution of eukaryotes.

Firming up the links between eukaryotes and Asgards would require more lab studies, and an example that grew more readily than the Japanese find. It took 12 years of trial and error to culture that first Asgard, and the second one from Slovenia wasn't much easier. "I didn't know how difficult it would be," says Schleper, who led the 7-year project.

With their long tentacles, the Asgard cells are fragile, a barrier to concentrating and growing them. Schleper's postdoc Thiago Rodrigues-Oliveira figured out culturing tricks that helped keep cells intact and tried different combinations of nourishment—including milk powder—until he hit a simple formula that encouraged growth. The resulting high concentrations enabled the team to create samples for cryo–electron tomography, a technique in which fast-frozen specimens are viewed at many angles by an electron microscope to develop a composite image. After solving a further challenge—picking out Asgard cells from other cells in the images—the team had a clear view of the microbes and their complex innards.

The new Asgard, which is different enough from those studied before to deserve its own genus with the tentative name *Lokiarchaeum ossiferum*, has thickenings and small bubbles along the tentacles. Its cell wall, too, is complex, with tiny lollipop structures sticking out, as if to sample the environment. "Overall, the cellular structures of [these cells] look like they come from another planet," Ettema says.

Its genome is larger and richer in eukaryotic genes than that of the other cultured Asgard, and it includes four genes for the protein actin, a key component of a eukaryotic cell's internal skeleton, Schleper's team reports. That skeleton extends throughout the cell and into the tentacles, and it varies from cell to cell, suggesting it's capable of being rearranged. "We show that the 'eukaryotic' cytoskeleton—which is crucial for eukaryotes—was an invention within archaea," Schleper says.

"This study further strengthens [the idea] that our ancestor is archaea," Imachi agrees. He and others now believe the most likely scenario for the emergence of eukaryotes is that an Asgard-like microbe enveloped other bacteria to create the complex eukaryotic cell, with its many organelles and compartments.

> **"Until recently, life's journey towards complexity was a blur."**
>
> **Masaru Nobu**, National Institute of Advanced Industrial Science and Technology

Others, however, aren't persuaded by this scenario. Sven Gould, an evolutionary cell biologist at Heinrich Heine University Düsseldorf, calculated from genomic data that Asgard archaea contributed very little to the first eukaryotes—as little as 0.3% of their predicted protein families. In a 10 November 2022 paper in *eLife*, he and his colleagues propose that complex features of the eukaryotic cell came not from an Asgard, but from the effects of stress on a much simpler ancestor after it absorbed bacteria. In this view, the stresses spurred the evolution of eukaryotic features such as the nucleus, the network of membranes and internal compartments called the Golgi apparatus, and even sexual reproduction.

Ettema agrees that the full story has yet to unfold. He notes that based on eDNA sampling by his group and others, the two cultured Asgards represent just a small subset of the group's diversity, and the true ancestor of eukaryotes may still be at large. So, he and others are trying to culture and characterize other Asgards.

"It will be exciting to see what other Asgard-like archaea are discovered and what they look like," Baum says. ■

U.S. BUDGET

Once banned, spending earmarks see resurgence

Lawmakers can't resist steering cash to universities and research projects back home

By **Jeffrey Mervis**

The $1.7 trillion spending package that President Joe Biden signed into law last week does more than fund the entire U.S. government for the first 9 months of this year. Senators and members of the House of Representatives from both parties also used it to funnel $15 billion to 7200 projects in their districts that federal funding agencies never requested. The projects include new research facilities and academic programs at hundreds of public colleges and universities.

That spending signals the robust resurgence of earmarks, the sometimes controversial—and until recently banned—practice in which legislators reward constituents using their constitutional authority over federal spending. The dollar amount and number of earmarks rose by half over last year, according to one count by *The New York Times*. The new total also tops levels seen before Congress banned the practice in 2010 after some notorious earmarks drew widespread ridicule—and figured in the conviction of one lawmaker for accepting bribes.

However, in early 2021 Congress removed the ban starting with the 2022 fiscal year. Proponents argued successfully that such directed funding addresses local and state needs, increases support for must-pass spending bills, and serves as a counterweight to the spending priorities of the executive branch.

Most earmarks do not relate to science; they typically fund transportation upgrades, new housing, or other forms of regional economic development. Still, research-related earmarks make many scientists uncomfortable. One complaint is

that those earmarks don't go through the merit-based competition used to allocate most federal research dollars. Science advocates also worry the growth in earmarks might shrink the pot available for competitively awarded grants.

Legislators acknowledge that the earmarking process can appear arbitrary. But even those who favor more funding for competitive research also seek earmarks, seeing them as another way to increase federal support for research.

"If there were a perfect science of how to go about [funding research], we would probably embrace that," says just-retired Representative Eddie Bernice Johnson (D–TX), outgoing chair of the House science committee. "But there is not. And so we try to use every opportunity to make sure that the importance of research is recognized." Johnson's earmarks in this year's appropriations bill include $2 million for Texas Tech University's Health Sciences Center in her Dallas district.

There's little research on how the outcomes of science-related earmarks compare with those produced by competitive awards. But a recent study of projects that legislators inserted into one Department of Energy program to fund hydrogen research found they generated as many publications and patents as projects chosen via merit review.

Clif Smart, president of Missouri State University (MSU), is one academic who doesn't see any downside to earmarks. The former teachers' college has a minuscule research portfolio—it ranked 321st among U.S. universities in 2021, according to data from the National Science Foundation—but has loftier ambitions. And when MSU needed a new science building, Smart turned to the state's senior U.S. senator, Republican Roy Blunt, who delivered $85 million in earmarks.

"Our science classrooms and labs are in buildings that are 50, 70, and 90 years old and that require more than $100 million to bring them up to today's standards," Smart says. "We would not have been able to do that without the directed spending that Senator Blunt championed for us."

Blunt, whose long tenure on the Senate appropriations committee has given him considerable influence, inserted earmarks worth more than $300 million into each of the last two federal spending bills. Those totals make him one of the top five earmarkers in Congress.

Senator Roy Blunt (R–MO, center) toured Missouri State University's aging science building with MSU President Clif Smart (right) before earmarking $85 million to renovate it.

If each of the 535 members of Congress secured as many earmarks as Blunt, this year's tally would have topped $160 billion. That's twice what nondefense agencies now spend annually on science. But Smart says the re-emergence of earmarks is not his concern.

"There probably should be some limits or constraints," he says. "But I'll let people smarter than me worry about what the caps ought to be. We've got 24,000 students [at MSU], 2500 of whom are majoring in science and another 8000 of whom take [science] classes. … My responsibility is making sure they have adequate facilities."

In reinstating earmarks, the Democrats who controlled Congress adopted rules designed to prevent the worst abuses. Members had to disclose every earmark they requested, and they couldn't fund for-profit entities or projects that would benefit them directly. A few legislators opted for even greater transparency, getting input from community leaders on what earmarks they should request. House members were also limited to 10 requests, although last year the cap was raised to 15.

Only a tiny fraction of the thousands of accounts that fund government activities are designated for earmarks. But their presence is significant at one science agency, the National Institute of Standards and Technology (NIST). This year, earmarks comprise more than half of the additional $103 million the agency received for its in-house research account and nearly one-quarter of NIST's overall budget of $1.6 billion.

Blunt and several other prolific earmarkers who didn't run for reelection in November 2022 went out with a bang. The reigning king, Senator Richard Shelby (R–AL), routed $762 million to his home state in the new spending bill. Shelby also pioneered the concept of federally funded endowments to hire and support faculty, sending a total of $100 million to the University of Alabama, Tuscaloosa, over the past 2 years. Another departing champion, Senator Patrick Leahy (D–VT), steered $212 million to his state, including a $30 million faculty endowment for the Honors College at the University of Vermont.

Blunt's total includes $61 million for research facilities at the state's flagship University of Missouri, Columbia, and $20 million for an innovation center at its Rolla campus. He has also funneled $6 million to MSU for an endowment to attract and retain science faculty.

MSU will miss Blunt's presence in Washington, D.C., Smart says. "My guess is that we will get substantially fewer directed expenditures," he predicts, noting that Missouri's two senators, including Blunt's successor, oppose the practice on the grounds that it wastes precious government resources.

But relatively few members of Congress have taken such a stance. House Republicans recently rejected by a three-to-one margin a proposal to reinstate the earmark ban after they assume control this week.

In Missouri, meanwhile, Blunt's generosity to MSU won't be forgotten. The current science building is named for Allen Temple, a pioneer in communications technology who chaired the school's science department for 40 years in the early 20th century. But last month the school's trustees renamed the new facility Roy Blunt Hall, and relegated recognition of Temple's long service to the university to a plaque in the atrium. ■

PHOTO: KEVIN WHITE/MISSOURI STATE UNIVERSITY

Found on the banks of an Idaho river, these ancient points resemble even older ones from Japan.

PHOTO: LOREN DAVIS

ARCHAEOLOGY

Ancient points suggest Asian roots for early American tools

Finds may support coastal route hypothesis for first settlers

By **Michael Price**

Thirteen razor-sharp projectile points found along a river in southwestern Idaho appear to represent the oldest evidence so far of toolmaking in the Americas—and may bolster the idea that the first people to reach the region migrated from Asia in boats along the coast of the Pacific Ocean.

Likely deposited into pits by a group of hunter-gatherers, the points were recently dated to between 16,000 and 15,600 years ago, according to a study in *Science Advances* last month. The tips are examples of "stemmed point technology," which allowed ancient people to fashion weapons from a wide range of available stone. Based on the objects' similarities to even older points found in East Asia, the blueprint for making the Idaho points may have come with the Americas' original settlers, the study authors propose.

A lot more work will need to be done to prove that, says Heather Smith, an archaeologist at Texas State University, but "it looks like a really interesting agenda to pursue."

The site where the points were unearthed a few years ago is on the banks of Idaho's Salmon River. The Nez Perce people, who have inhabited the region for thousands of years, refer to it as Nipéhe, for an ancient village there. In English, it became known as Cooper's Ferry.

Sixteen thousand years ago, the river sat in an ice-free corridor. At the time, an overland route into the North American continent from the Bering Strait would have been blocked by massive ice sheets. But some researchers have proposed that the earliest migrants from Siberia boated along the shore of an ice-covered Bering Strait. "If you're coming south along the Pacific coastline entering North America … the first major left-hand turn south of the ice is the Columbia River, and if you head upstream, you can get to Cooper's Ferry," says Loren Davis, an Oregon State University, Corvallis, archaeologist who led the new study.

Devastating floods and avalanches destroyed or buried nearby valleys, but Cooper's Ferry was relatively unscathed, he says. "As far as we can tell, people early on decided this was a really great place to live, and they kept coming back over and over and over again." That tracks with Nez Perce history, says Nakia Williamson-Cloud, cultural resources program director for the tribe, on whose lands the artifact-filled site sits. Stories passed down over thousands of years tell of a young couple founding the village after a catastrophic flood destroyed their previous home across the river.

Davis began working at Cooper's Ferry in 1997 as a graduate student and never left. In 2019, he and colleagues published a paper in *Science* that included tentative radiocarbon dates of between 16,560 and 15,280 years for bits of animal bone and charcoal excavated in collaboration with the Nez Perce Tribe, making it one of the earliest known human-occupied sites on the continent.

For the new work, Davis's team—including interns from the Nez Perce Tribe—went 25 meters upriver from previous digs and unearthed three cylindrical pits that contained the spear points and hundreds of bits of bone, likely from animals, that an outside lab radiocarbon dated to between about 16,000 and 15,600 years ago, firming up the dates reported in the earlier study.

Smith says the new analysis brings "needed rigor" to the previous Cooper's Ferry dates. But Ben Potter, an archaeologist at the University of Alaska, Fairbanks, remains unconvinced, arguing the artifacts from the pits are too jumbled to conclusively link them to any of the animal bone dates. "Their precise age remains unclear, in my opinion."

Although no genetic evidence connects the ancient toolmakers to modern Nez Perce people, Williamson-Cloud says he believes "these are truly our ancestors. … It's a place where our lineage came from—people who are alive today."

Davis and the study's other authors, who include Japanese and Chinese archaeologists, think there's a good case to be made that the first migrants from Asia brought with them the rough-and-ready stemmed point technology, with fluted ends that were wedged into spear tips. The points at Cooper's Ferry, they say, most closely resemble projectile points made by people who lived near modern-day Hokkaido, Japan, some 20,000 years ago.

Genetic studies show these people were not ancestors of modern Native Americans, but Davis believes they may have passed technological traditions on to other Asian groups that did migrate through northeastern Siberia and into the Americas. "[Those travelers] didn't invent this stuff when they got to the Americas," he says. "When they left northeast Asia, they had a whole set of technological ideas in their minds."

Davis's scenario makes sense to Matthew Des Lauriers, an archaeologist at California State University, San Bernardino, who studies stone tool technologies. David Meltzer, an archaeologist at Southern Methodist University, remains skeptical. He says the similarities between the two regions' stemmed points appear generic. Finding more evidence of human artifacts at sites in between Japan and the U.S. Pacific Northwest would help make the authors' case, Meltzer adds, but "detecting actual links between populations so distant in space and time can only be done reliably with ancient genomics." ■

A handheld air sampler keeps watch for airborne viruses at a Cambodian market. The device can also be loaded onto a drone or toy car to reach inaccessible spots.

INFECTIOUS DISEASES

Virus hunters test new surveillance tools

Ropes, drones, insects, and dust cloths could make monitoring faster, safer, and cheaper

By **Jon Cohen**, *in Phnom Penh, Cambodia*

On a Friday morning in September last year, Erik Karlsson visited the sprawling Orussey market here, where vendors hawked pots and pans, phone cords and radios, hats and dresses—and myriad types of Southeast Asian food. Dozens of orange-colored, whole roasted pigs hung on hooks, crabs the size of two fists filled buckets, and stacked fruit and dried fish formed mountains on tables.

Karlsson had come for the live poultry, but not because he was planning a dinner. As an epidemiologist at Cambodia's Pasteur Institute, he was hunting for potentially dangerous pathogens, both known and unknown. He had nothing with him but a cellphone in one hand and, in the other, a device that resembles a portable credit card reader, which he held close to men who were slaughtering, boiling, and defeathering chickens. The AeroCollect, as it's called, has a chip that uses an electric field to suck in and trap air in microscopic chambers. Later, back at his lab, Karlsson flushed each chamber's content with a water rinse and ran a polymerase chain reaction (PCR) to amplify any viral RNA in the air.

Karlsson is one of many researchers looking for simpler, faster, cheaper, and safer ways to find viruses—including the ones that could cause the next pandemic. If the AeroCollect works well, he won't need to don masks and gloves and get permission to draw chickens' blood and swab their behinds. Karlsson also plans to test flying the device on a drone into bat caves—which are known viral hotbeds—so researchers don't have to enter them.

Other scientists are tapping advances in nanotechnology, sequencing devices, artificial intelligence, and robotics to improve virus hunting and provide a more timely, deeper understanding of potential threats. Programs to monitor SARS-CoV-2's spread and evolution using wastewater (*Science*, 11 March 2022, p. 1100) have shown how new methods can help.

Many of the projects, including Karlsson's, are at the proof-of-concept stage, but Maria Van Kerkhove, who heads the emerging diseases and zoonoses unit at the World Health Organization, welcomes them. "I'm really excited about the stuff Erik is doing," says Van Kerkhove, who sampled birds at Orussey market herself more than a decade ago as part of a Ph.D. project on bird flu. "Air sampling is leapfrogging in technology."

But new surveillance technologies must not lead to what Van Kerkhove calls "a stamp collecting exercise": simply identifying the hundreds of thousands of viruses in wild and captive animals. "Detection is only one element," she says. "Does the virus have the ability to spill over from animals into humans? And if so, what does that cause? Is it a dead end or does it have an epidemic outbreak risk?" Just such a criticism has dogged the Global Virome Project, an ambitious scheme that promises to find the vast majority of pathogens that threaten humans but has struggled for many years to find the funding to launch (*Science*, 23 February 2018, p. 872).

To properly gauge the risk of outbreaks, researchers have to conduct repeated surveillance for viruses in spillover hot spots such as live-animal markets, farms, and bat caves, and at the clinics near them, says epidemiologist Christine Johnson, who directs the EpiCenter for Disease Dynamics at the University of California (UC), Davis. "I don't think we just want to do a snapshot in time," she says. "It has to be fundamentally more of a long-term, watching things evolve to really gain any true insight or understanding."

Still, most virus hunters agree a bit of ingenuity and creativity can make identifying threats a lot easier. Here are some of the avenues they are exploring.

JOHNSON'S TEAM is giving wild monkeys something to chew on—literally. It is exploring whether scientists can sample the animals' saliva by giving them treats on ropes.

PHOTO: J. COHEN/*SCIENCE*

The project grew out of the U.S. government's PREDICT, a 30-country effort launched in 2009—and run by UC Davis—to identify viral threats. Sampling wild primates traditionally requires trapping and anesthetizing the animals. That raises ethical and practical issues, however, so some scientists have simply collected feces or urine instead. But those approaches have drawbacks as well: They often require direct observation or accessing the animals' nests in trees, and neither is ideal for viruses that primarily shed through the mouth and nose.

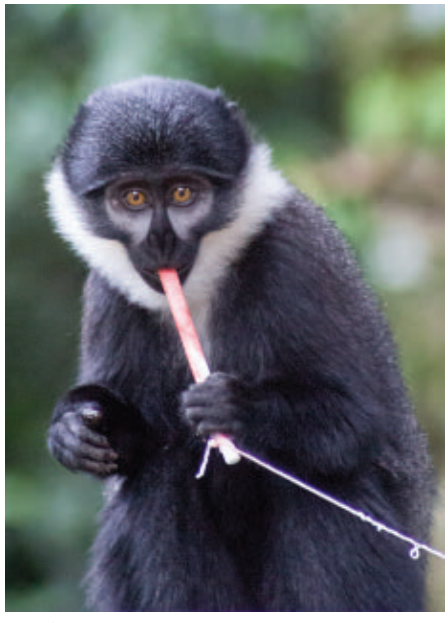
These limits led Johnson and veterinarian Tierra Smiley Evans to test ropes laced with treats such as strawberry jam or banana baby food that monkeys chew on, leaving behind gobs of saliva. In proof-of-principle studies in Uganda and Nepal, they found simian foamy virus, parainfluenza 3, enteroviruses, and herpesviruses. The group has also collected plants chewed and discarded by Uganda's endangered mountain gorillas, most recently to assess whether people were infecting them with human herpesvirus-1 (HSV-1), which has sickened and killed captive lowland gorillas. There were no signs of HSV-1 in the plant discards from 294 gorillas, the team reported in October 2022 in the *American Journal of Primatology*, but they did contain gorilla-specific herpesviruses.

UC Davis has more futuristic projects in the works. A team is exploring whether a sandwich-size air sampler can identify active viral infections in humans from patterns of volatile organic compounds (VOCs) in their breath. Several respiratory diseases, including influenza, tuberculosis, and COVID-19, come with specific molecular breath signatures, researchers have found. In a December 2022 paper in *Communications Medicine*, a UC Davis team led by Mitchell McCartney and Cristina Davis even reported distinctive breath markers during the spread of the Delta and Omicron variants. The group has also tested the air sampler's ability to detect disease signatures while attached to a person's hip or flown on a drone.

The "breath biopsy" field is still in its infancy, despite a surge of interest during the pandemic in room air samplers and breathalyzerlike devices. (The Netherlands used SARS-CoV-2 breathalyzers for widespread COVID-19 testing, but officials soon deemed them unreliable.) In their attempt to make "breath the new blood" for diagnoses, Davis and an international group of colleagues last year launched the Human Breath Atlas, an effort to conduct large-scale investigations into the rare VOCs people exhale when sick.

IN THE NETHERLANDS, veterinarian Marion Koopmans and colleagues have been placing electrostatic cloths in chicken barns to collect dust that can be analyzed for viruses. "There's remarkably little high-quality information on viromes of farmed animals," she says. "We're exploring ways to make surveillance more accessible."

There are 22 billion chickens in the world, according to a 2017 estimate, and they harbor dangerous viruses—H5N1 influenza is the most famous one—that occasionally spill over into humans and might cause the next pandemic. Earlier this year, Koopmans and colleagues repeatedly visited three farms and compared dust snagged by the cloths with chicken feces scooped up by hand, the usual way to look for avian viruses. They didn't simply test the samples for known viruses, but used metagenomics to sequence any genetic material in the sample. Samples from both sources yielded similar results—viruses from four families—confirming the "electrostatic dustfall collectors" are a good alternative to sampling poop, the group reported in *Scientific Reports* last year.



Scientists analyzed viruses left on this rope after a monkey ate an attached treat.

PHOTO: TIERRA SMILEY EVANS

Koopmans hopes the technique can help scientists forecast where trouble is brewing. "Right now, the way our surveillance works is really very reactive," she says. "How can we move surveillance forward so we pick up things before you see human disease?"

SOME VIRUS HUNTERS want to send in an army—of ants. Voracious army ants prey on many insect and vertebrate species and can travel about 1 kilometer a day, often in and out of dense forests that researchers cannot easily enter. A team led by plant virologist Philippe Roumagnac at the University of Montpellier wondered whether the ants pick up the pathogens infecting their meals along the way. In Gabon, Roumagnac and his colleagues captured 209 army ants from 29 different colonies of the genus *Dorylus*, ground them up, and conducted a metagenomic survey of the DNA inside them. Sequences from 157 different viral genera popped up, the group reveals in a preprint posted last month on bioRxiv. Metagenomic analyses can also reveal what species the ants ate, helping tie specific viruses in a sample to what they may have infected. Using ants avoids focusing on animals that humans think pose the greatest risks, Roumagnac says: "They have no sampling bias."

To validate the method, the team plans to sample bat guano in Gabon for viruses and compare them with what they find in army ants feeding on the same guano. Roumagnac is optimistic: "Army ants could be one of the key players that could give clues about the bigger picture of the ecology and evolution of viruses in a single ecosystem."

KARLSSON'S PHNOM PENH OFFICE is filled with what he calls "toys" that might expand the AeroCollect's surveillance powers. A drone could be used to sample bat caves—even ones that only rock climbers can now reach. Karlsson recently purchased a remote-controlled, pint-size car that can hold the device—he plans to drive it around farms to collect samples from a distance. Battery-powered, downsized PCR devices and portable sequencing machines will accelerate the analysis. "We want to bring the lab to the field," Karlsson says. "If we're at an outbreak, we can get ahead of it faster."

The Pasteur team that Karlsson joined 5 years ago has a long history of working to improve viral surveillance in Cambodia, focusing on how bird flu evolves and spreads to humans. The team also understands the challenge of translating its findings into actions that protect people—a challenge underscored by PREDICT, which, despite an investment of more than $200 million, did not identify the threat of SARS-CoV-2. Veasna Duong, who heads the Pasteur Institute's virology division, adds that preventive measures require political will and acceptance by local people who often don't accurately perceive the risk.

That's no reason not to try, Karlsson says. "Can we prevent pandemics?" he asks. "I think we can be better prepared for the next one, that's for sure." ■

Reporting for this story was supported by a grant from the Alfred P. Sloan Foundation.

FEATURES

A RIVER'S PULSE

In the Brazilian Amazon, Indigenous people and researchers join hands to monitor the impacts of a controversial dam

By **Sofia Moutinho**; *Photography by* **Dado Galdieri**

Soon after sunrise one warm day in September 2022, 26-year-old Josiel Pereira Juruna boards a small motorboat and sets out on the emerald-green waters of the Xingu River in the Brazilian Amazon. Accompanying him are biologist Cristiane Carneiro and Pedrinho Viana, a fellow fisherman from their village of Muratu in the Paquiçamba Indigenous Reserve in northern Brazil's Pará state. After a short ride, Viana hauls in a gillnet set out in a creek the night before.

This story was produced with support from the Rainforest Journalism Fund in partnership with the Pulitzer Center.

He pulls a disk-shaped fish with bright gray scales and a yellowish belly from the net and hands it to Josiel, who hangs it from a portable scale. "One hundred and fifty grams," he declares, then presses a ruler against the animal, known as a big-eyed pacu. "Fifteen centimeters," he says, as Carneiro takes notes.

It's a ritual of weighing and measuring that Josiel has performed nearly daily for the past 3 years to monitor the river's fish stocks. Of all fish in this stretch of the Xingu, the seven species of pacu are the most important for his community, the Juruna, who rely on fishing for food and income. Known as vegetarian piranhas, pacu can reach up to 1 meter long. But they are dwindling. In November and December 2014, the best fishing months of that year, fishermen in Muratu caught a total of 770 kilograms of pacu. Over the same months in 2021, that number dropped to 175 kilograms.

"Now we only catch very skinny and small pacu, almost without fat," Josiel says with disappointment as he gestures to the specimen in his hand.

The likely cause lies 30 kilometers upstream of Muratu: the largest hydroelectric project in the Amazon basin, a complex of dams, reservoirs, and power stations known as Belo Monte. The project was originally slated for construction in 1975, but years of protests from Indigenous communities and a lack of investment stalled it. In the 2000s, after major electricity blackouts in

Fishermen from the Indigenous village of Muratu catch a pacu, a fish that provides sustenance and income along the Xingu River in Brazil. Indigenous researchers are tracking pacu numbers and size (left) to monitor the effects of a giant dam just upstream.

PHOTO: DADO GALDIERI/HILAEA MEDIA

the country, President Luiz Inácio "Lula" da Silva (recently reelected) pushed for the project, despite international opposition from environmentalists and scientists, and it began operating in 2016.

The main dam, called Pimental, had a dramatic impact on the river, creating a 359-square-kilometer reservoir and diverting much of the Xingu's flow northeast through a 17-kilometer canal to a secondary reservoir and hydroelectric station. Downstream, along the 130-kilometer river stretch called the Big Bend of the Xingu, the diversion reduced the river's flow by up to 80%. It also interrupted the river's annual cycle of flooding, crucial to its rich biodiversity. The Juruna call 2016, the year the main dam was completed, "the end of the world."

The impacts were evident right away, when an abrupt release of water from the dam killed 16.2 tons of fish. The Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) imposed a fine of 35.3 million reais ($6.6 million) on Norte Energia, the company that runs the hydroelectric project. Josiel and other people from the Paquiçamba reserve were not surprised at the havoc. "The elderly in our villages knew we would have this kind of problem with the dam," Josiel says. "We knew the fish would suffer with a river flow that is no longer natural."

They had already asked researchers to help them document the changes. In 2013, with the support of a local nongovernmental organization known as the Socio-Environmental Institute, the Juruna established a partnership with an informal network of scientists from public universities across Brazil. Among them was Carneiro, who was doing a Ph.D. in ecology at the local Federal University of Pará (UFPA), Belém. Together, the Juruna and the scientists have been recording change in the abundance, habits, and size of the fish and turtles living in the changed river.

For Philip Fearnside, a biologist at the National Institute of Amazonian Research (INPA), the collaboration "is essential to show the huge impact instead of simply accepting it." Fearnside, who isn't involved in the collaboration but has been studying dams in the Amazon for more than 30 years, says partnerships between Indigenous peo-

The Belo Monte hydroelectric project is the Amazon basin's largest. The main dam (above), called Pimental, diverts much of the Xingu River into a secondary reservoir, where it drives turbines.

A woman from the Juruna community applies ceremonial paint to biologist Janice Muriel Cunha (right) during a September 2022 meeting to discuss research ties between Indigenous people and scientists.

ple and researchers are critical in Brazil, where government approvals to build and operate dams are often "a token gesture," and companies "can go ahead and build the dam anyway," regardless of impact.

In a statement sent to *Science*, Norte Energia said changes in Xingu fish populations are within the range of what was predicted in the environmental impact studies conducted before the dam's construction. But the full scope of those impacts is emerging in the data the Juruna and their scientist collaborators have collected. Their findings have been published in peer-reviewed journals and used to fight for reparations and for the dam's operators to restore some of the river's natural flow.

Now that Lula is once again president, after campaigning on promises of a strong environmental agenda and plans to create an Indigenous Ministry, the Indigenous people and their scientist allies hope their findings will lead to permanent changes at Belo Monte—and in the hundreds of other dams being planned or built in the Amazon. "I think things will get better for Indigenous peoples during this government and maybe more water will be released to the Big Bend," Josiel says. "That is what I hope for."

PHOTOS: DADO GALDIERI/HILAEA MEDIA

Threatened fish nurseries

The Juruna mapped more than 30 fish breeding sites in flooded forests and islands along the Big Bend of the Xingu, a 130-kilometer river stretch with unique biodiversity and surrounded by Indigenous villages. These seasonally flooded areas hosted multiple species of fish and turtles but have been mostly dry since the construction of the Pimental Dam.

GRAPHIC: K. FRANKLIN/*SCIENCE*

BEFORE THE DAM, the Juruna—and the fish—could count on the river's seasonal rhythm. During the rainy Amazonian winter, from December to May, as much as 20,000 cubic meters per second of water surged down the river, overflowing its several channels and spilling onto islands and adjacent forests. In the summer the flow falls as low as 2000 cubic meters per second. Over millennia, this flood pattern has shaped the river's landscape and species.

Now, the dam has cut the winter flow by more than half. Worse is to come: a new flow regime that in alternate years will reduce the maximum winter flow by another 50%, to one-fifth the natural level. (IBAMA had approved the plan, but until recently courts had blocked it.) Because of the water diversion, "The cycle of the flood pulse, the most basic characteristic of that ecosystem, was completely altered," says Camila Ribas, an INPA biologist who collaborates with the Indigenous peoples. "If you change this, you change all the relationships within this system and may permanently destroy it."

The Juruna and their scientist partners have been monitoring the consequences. Outside researchers trained 12 people from the Juruna community to conduct the work. Every day since the project began in 2013, one of them is tasked with recording the weight and length of individual fish, as well as the total amount of fish caught in the reserve.

More recently, the Juruna also started to investigate how the change in the river flow has affected fish breeding sites in the *igapós* (from the Guarani language for "root forests"), the seasonally flooded forests on the river margins and islands. For multiple species of fish and turtles, these forests are vital habitats. Pacu, *matrinxã*, *curimatá*, and other local fish feast on ripe fruit that falls into the water from the trees, a bounty that includes the tiny red cherrylike *sarão* (*Myrciaria dubia*), the yellow plum-shaped *cajá* (*Spondias mombin*), and the round seringa, the fruit of the rubber tree (*Hevea brasiliensis*).

In 2019, Josiel and a team of researchers from UFPA installed flood gauges on Zé Maria, a forested island 4 kilometers from Muratu village. Every November before the dam was built, the rain-swollen Xingu would slowly advance over the island's sandy beaches, reaching the top of its trees in April. The island remained submerged in up to 30 meters of water for the next 6 months, bringing an abundance of fish to feed and spawn. Now, even in March, Josiel can stand on the island's leaf-littered ground. The instruments have never measured more than 2 meters of water.

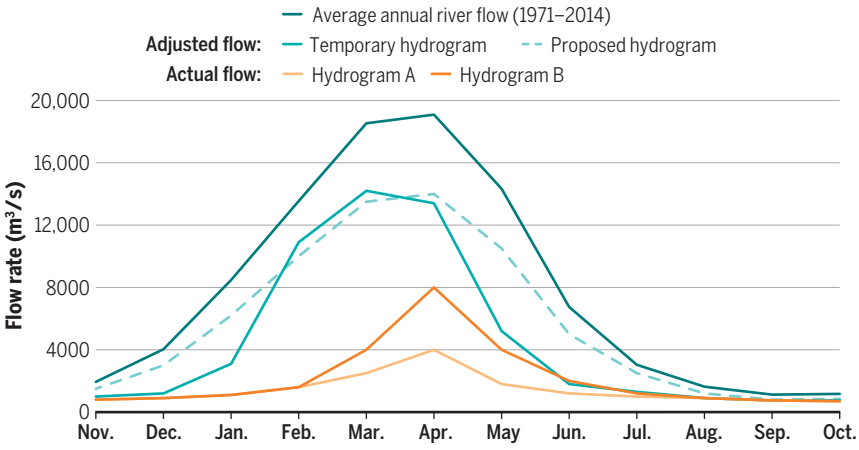
Up to 70% of the Big Bend's seasonally flooded forests no longer flood under the current water regime, according to data from Norte Energia. Other Amazon dams likely have similar impacts. Although Brazil requires environmental assessments for dam projects, the laws don't specifically mention flooded habitats such as the *igapós*. "Our legislation has not kept pace with the scientific knowledge," says Andre Sawakuchi, a geologist at the University of São Paulo's main campus who is involved in the effort with the Juruna. "*Igapós* are forests inside the river. If you lose them, you cannot just replant. You cannot replant a river."

Without a place to spawn, fish have been found with dry eggs inside their bodies "like dead babies inside a woman's belly," Josiel says. The fish have also lost an important source of food. "The fruits now fall onto dry ground, and the fish cannot eat," Josiel explains. "It is very sad."

The resulting decline of the Big Bend fisheries has hit the Juruna hard. "Before

A weakened pulse

In natural conditions, the Xingu River has a high volume during the Amazonian "winter" (December to May) and a low volume in the summer. But the construction of the Pimental Dam drastically reduced the water flow. (Hydrograms A and B represent alternating flow regimes.) In 2019, the government temporarily forced the dam to release more water. Now, scientists and the Juruna community have proposed increasing the flow to minimize environmental damage in the river.



In winter, the swollen Xingu River normally floods Zé Maria island to a depth of tens of meters. A flood gauge placed by Indigenous researchers recorded last winter's much-reduced flood.

[the dam], we could fill three coolers of 160 liters with fish in a week," remembers Maria das Graças, a Juruna fisherwoman and Viana's wife. She and her husband raised their three children on fish from the Xingu. What the family didn't consume, they used to sell. "Now, we go out fishing for 10 days and can barely fill two coolers."

The data collected by the Juruna have quantified the change. Sent to researchers at UFPA for further analysis, the findings have led to academic papers on Xingu ecosystems and changes in Indigenous ways of life, and have gone into monthly reports presented in open meetings in the villages.

A separate study published in *Science of the Total Environment* in September 2022 confirms the Juruna's observations. The research, partially funded by Norte Energia and conducted with data the company was required to collect, found a 29% decline in the number of species and a 9% drop in the abundance of all fish in the Big Bend. The decline was sharpest for pacu. The results are just the beginning of a "tragedy" that will worsen, says Kirk Winemiller, an ichthyologist at Texas A&M University, College Station, and an author of the study, because "we are only looking at the very early signs" of the dam's impact.

Winemiller thinks some species may face extinction. The catfish acari-zebra (*Hypancistrus zebra*), a match-size fish with black and white stripes only found within the Xingu's Big Bend, is one of his biggest worries. Illegal fishing for the aquarium trade had already depleted it, and in 2022 the International Union for Conservation of Nature declared it critically endangered because of the dam.

THE COLLABORATION'S FINDINGS are already having an impact. The data are often sent to IBAMA—and the Brazilian Federal Public Prosecutor's Office, which has filed several lawsuits against Norte Energia. Some accuse the company of "ethnocide," the crime of destroying Indigenous cultures, because of the severe impacts of the dam on their lives.

In its statement to *Science*, Norte Energia said it has compensated the communities. New fish farming tanks installed in Indigenous villages have provided 660,000 reais ($124,000) in income to local families since 2019, the company says. In November 2022, the company also agreed to pay an undisclosed sum in reparations to about 2000 local fishermen. IBAMA had ordered this compensation as a condition for renewing the dam's operating license.

For the Juruna, no money or mitigation project can substitute for the altered pulse of the river and their own lives. "Belo Monte took our river, and it was like draining our blood out," says Giliarde Juruna, Muratu's chief. "Today, we are adapting to this new reality of drought. But we don't know how our future will be without the river."

The collaboration hopes its data can be used to push for more water to be released from the dam. In 2019, based on the Juruna's data, IBAMA deemed the dam's water flow insufficient and forced Norte Energia to temporarily increase it, reducing the amount of water delivered to the power turbines. IBAMA also asked the company to conduct further impact studies. In 2021, however, then-President Jair Bolsonaro's administration reversed the ruling and allowed Norte Energia to reinstate the original, reduced water flux. The agency did require the company to invest 157 million reais ($29 million) in environmental and social mitigation actions, such as replanting trees.

In August 2022, the Juruna and their scientist partners sent IBAMA another proposed water regime that would increase the volume released from the dam by up to 68%. The increased flow would restore flooding to 32 mapped *piracemas*, or fish breeding sites, like those on the island of Zé Maria. The plan would also do away with the abrupt daily changes in flow that the river experiences now, like the one for which Norte Energia was fined in 2016. "The water needs to go up and down gradually to guarantee that the fish juveniles can develop," says Janice Muriel Cunha, an

CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) BRAZILIAN NATIONAL AGENCY OF WATERS, ANALYZED BY ANDRE SAWAKUCHI/FEDERAL UNIVERSITY OF SÃO PAULO; (PHOTO) DADO GALDIERI/HILAEA MEDIA

The acari-zebra, unique to the Big Bend of the Xingu River, already faced pressure from collectors supplying the aquarium trade. The dam and its impact on the river could push the fish toward extinction.

PHOTO: DADO GALDIERI/HILAEA MEDIA

ichthyologist at UFPA. "They need at least 3 months to grow in the calm waters of the flooded forests, and then go back to the main stream."

IBAMA is now evaluating the suggestion and should answer soon. But approval is unlikely, the researchers admit. More water down the river means less flowing to the turbines, and less electricity produced by a complex that is already falling short of expectations. Because of drought and poor planning, Belo Monte only generates on average 4571 megawatts of power, less than half its installed capacity. "Based on the history so far, it is difficult to believe that Norte Energia would reduce energy production to increase the water volume in the Big Bend," Sawakuchi says.

THE JURUNA HOPE things will take a better turn with Lula. It's unclear how the politician, who said in a recent interview that he would approve the construction of Belo Monte again, will follow through on his campaign promises on the environment. Biviany Rojas, who coordinates the program that supports the Juruna's monitoring project at the Socio-Environmental Institute, expects that under the new government, IBAMA will recover the autonomy it had lost during Bolsonaro's administration to make decisions based on technical information rather than on politics. "This is an opportunity for Lula to redeem himself with the Xingu people," she says.

For his part, Josiel is confident there will be more dialogue between his people and the new government. But he remains concerned about the future, and not just for his people. There are about 150 other dams in the Amazon basin already, and 350 more are planned—including another on the Xingu, upstream from Belo Monte.

Although he only has a high school degree so far, Josiel has been thinking about becoming a biologist. The challenges to achieving that goal are multiple, but he thinks science can give a voice and memory to the knowledge of his people, passed down from generation to generation, including what he learned from his grandfather about the river and how to live from it. "It is a way to show future generations of our community how the river was before, how abundant the fish was," he says. "And how it will never be again." ■

NEXTGEN VOICES

The fruits of failure

We asked young scientists: **Are you grateful for a setback that you experienced? What mistake did you make, and what good came from it?** In the responses below, they share the discoveries and revelations that made their would-be regrets worthwhile. Follow NextGen Voices on Twitter with hashtag #NextGenSci. **—Jennifer Sills**

Interdisciplinary collaboration

After many unsuccessful attempts to produce diffraction-quality protein crystals, I finally asked others for help. By doing so, I changed my routine from solo lab work to collaboration, which gave me a broader understanding of the challenges faced by experts, peers, and junior scientists. Viewing mentorship and science funding from the lens of a struggling scientist led me to a new career. I became a grants manager at a funding body, where I was responsible for science communication, grant writing workshops, collaboration, and networking events for scientists. As a science administrator, I could enhance the quality of science through funding, building capacity for addressing the "behind the scenes" matters of research. If I had not faced failure working in silos in a lab, I would never have discovered my passion.

Garima Singh
Mott MacDonald, Noida, Uttar Pradesh 201301, India. Email: garima.singh@mottmac.com

Two years ago, I was trying to construct mathematical models for predicting the risk of maternal inheritance of pathogenic mtDNA mutations. As a medical student, I seldom receive mathematics training. After struggling for a month, I turned to a friend studying information science for help. He solved the functions easily, making me realize the importance of interdisciplinary collaboration. Last year, we established an organization dedicated to promoting the integration of medicine and other disciplines, including bioengineering, computer technology, and artificial intelligence.

Ning Zhang
First School of Clinical Medicine, Anhui Medical University, Hefei, Anhui 230032, China.
Email: ningzhangbox@163.com

Writing opportunities

I'm currently taking 2 extra years of college because I wanted to keep up an adequate academic performance in classes and extracurricular research projects as well as prioritize my physical and mental health. I worried that the delay would mean missed opportunities, but I now have more time to participate in scientific events, develop friendships, and stand out in classes. When I found out that I was accepted to an exciting writing internship, I realized that saying no to certain options is saying yes to others. I may not graduate in the expected time for most people, but I'm graduating in the right time for me, while doing something I love: communicating science in an accessible, entertaining, and useful way.

Camila Fonseca Amorim da Silva
Olá, Ciência!, University of São Paulo, São Paulo 03828-000, Brazil.
Email: camilafonsecaamorim@usp.br

I signed up for a course at a local museum, hoping for a research experience. On my first day, however, I realized that I had accidentally signed up for an intensive writing seminar. Little did I know that the course would expose me to an area of science that would prove invaluable in my career.

ILLUSTRATION: ROBERT NEUBECKER

The seminar taught me how storytelling is central to sharing scientific discoveries. Were it not for that seminar, I would never have undertaken journalism internships, developed a passion for science writing, or become as thoughtful a scientist.

Divyansh Agarwal
Massachusetts General Hospital, Harvard Medical School, Boston, MA 02108, USA.
Twitter: @divyansh_aga

Reaffirmed values

Several years after publishing our research findings, we realized that we had made some errors. We promptly contacted the journal, which decided to retract the manuscript. Because we had been transparent, we were allowed to submit a corrected version, which was later accepted. Although retractions are often seen as academic failures, going through the process reminded me that integrity and honesty are the cornerstones of our research and made me feel more committed to our values as scientists.

Salam Salloum-Asfar
Qatar Biomedical Research Institute, Hamad Bin Khalifa University, Qatar.
Twitter: @Dr_SalamSalloum

With only a bachelor's degree in physics and chemistry and a sense of purpose, I tried to create a hydroponic system that would lower the price margins in my country and provide fresh, organic, and affordable vegetables. The journey was rough, and the start-up suffered heavy financial losses. However, my efforts led to more fresh organic produce in Trinidad's supermarkets, an award for ingenuity, and two publications. Those successes inspired me to venture into medical research, where I found opportunities that brought me closer to my ultimate goal of contributing meaningfully to humanity.

Stephanie Mohammed
Department of Physics, Faculty of Science and Technology, The University of the West Indies, St. Augustine Campus, Trinidad, St. Augustine, Trinidad and Tobago.
Email: stephanie.mohammed@sta.uwi.edu

Passion for teaching

As a first-generation undergraduate, I was excited to be accepted by a PhD program in zoology immediately after I graduated. However, I didn't take the time to consider my career goals, and after a year of frustration, I began to doubt whether I was cut out for grad school. Because my solace from research struggles was instructing undergraduate biology, a mentor recommended looking into discipline-based education research programs. Only after I was accepted into a doctoral program in biological education did I realize how wonderful research could be.

Ashley Barbara Heim
Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA.
Email: abh229@cornell.edu

I planned to major in chemistry and then work in the lab of some industry or research facility. However, after failing to receive a scholarship I applied for, I had to find another source of income. I had never thought of myself as a teacher before, but I took a job preparing students for their final exams. Suddenly, I realized that teaching was my vocation and quickly updated my academic plan. Now an assistant professor at the university level, I aim to help students follow this path, one I may never have found if I had secured that scholarship.

Wagner Eduardo Richter
Department of Chemistry, Federal University of Technology–Paraná, Ponta Grossa, Paraná, Brazil.
Email: richter@utfpr.edu.br

Improved procedures

After my manuscript was published, I realized, to my horror, that we had used identical reference protein images in two separate results. I checked my original data archives and discovered how difficult it was to find the raw images. After providing the right pictures to the journal, I immediately asked a computer expert to help me develop a better data storage system. Now I upload my raw data daily, all processed data is organized and saved, and all files are backed up weekly. Publishing a mistake taught me the importance of proper data management.

Yongsheng Ji
Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230026, China. Email: jiys2020@ustc.edu.cn

Once, after a chemistry experiment, I poured liquid waste containing ammonia directly into the toilet. The strong odor it produced did not go unnoticed by my supervisor. Remembering the smell and the subsequent criticism I received has motivated me to take extreme care when conducting and cleaning up experiments ever since, possibly saving me from accidents and injuries.

Yuan Zhi
School of Economics, Guizhou University, Guiyang, Guizhou 550025, China. Email: yzhi@gzu.edu.cn

New research avenues

One night during medical school, my last task was to take out two western blots for two different experiments and place them in antibodies specific against different proteins. The next day, I realized that in my exhaustion I had swapped the blots. Because of this silly mistake, I stumbled upon a novel modification of my protein of interest! The solution to a 2-year mystery came into focus. This serendipitous finding led me to pursue a full-time PhD course to explore a new avenue of research.

Bryce W. Q. Tan
Department of Medicine, National University Hospital, Singapore 119228, Singapore.
Email: bryce_tan@nuhs.edu.sg

After a tough semester in my undergraduate physics education, I had not completed enough courses to be eligible to continue with the mandatory undergraduate thesis project the same year. I was devastated. Soon afterward, I discovered a new field of research that sparked my interest. The field—biophysics—was outside the traditional topics offered at my institution, but I found a research group and convinced the institution to let me do a biophysics project. Had I not failed and postponed my thesis, I would probably not be working in biophysics today.

Cathrine Bergh
KTH Royal Institute of Technology, 11423 Stockholm, Sweden. Email: cabergh@kth.se

Supportive relationships

Because I initially overlooked the acceptance email for a summer program, I missed the window to submit my preferences. I was auto-matched to a venom lab, where I learned to my consternation that I would be hand collecting spiders at midnight from a forest. However, this unexpected topic came with enthusiastic mentors. No snake, turtle, or scorpion went without a photo, and the group took lengthy detours just to appreciate the scenery. I was new to research, and good mentorship and memorable experiences outweighed all else. The experience fueled my interest in pursuing science as a career.

Jackson Ross Powell
Vagelos Molecular Life Sciences Program, University of Pennsylvania, Philadelphia, PA 19104, USA. Email: jrp24@sas.upenn.edu

In high school, I found a welcoming mentor who catalyzed my love for research. I asked him if I could work in his lab during the summer, and he jumped through hoops to include me. When I told him later that I had applied to other programs as well and had accepted another offer, he was understandably livid. By taking him for granted and failing to understand the logistics of lab staffing, I ruined my relationship with him. Today, remembering my mistake constantly reminds me to appreciate and communicate with my mentors.

Sai Sarnala
Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA.
Email: ssarnala@sas.upenn.edu

10.1126/science.adg1443

A University of California San Diego researcher collects wastewater on campus for SARS-CoV-2 analysis and sequencing.

EPIDEMIOLOGY

Wastewater surveillance for public health

Wastewater contains information on pathogen spread, evolution, and outbreak risk

By **Joshua I. Levy[1], Kristian G. Andersen[1], Rob Knight[2,3,4], Smruthi Karthikeyan[2]**

Dating back to the origins of modern epidemiology, wastewater surveillance has predominantly been used to track pathogens spread by fecal-oral transmission such as those that cause cholera and polio. However, more than just these "enteric" pathogens are shed via the gut, as highlighted by the success of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) wastewater surveillance (*1*, *2*), recent work on tracking influenza virus (*3*) and monkeypox virus (*4*), and observations of extensive pathogen diversity in stool (*5*, *6*). Wastewater is now a core component of infectious disease monitoring, providing a variant-specific, community-representative picture of public health trends that captures previously undetected spread and pathogen transmission links. Building on recent laboratory and analytical advances to identify the diverse pathogens present in sewage will be essential to ongoing efforts to understand disease risks and will transform infectious disease surveillance.

Pathogen surveillance typically involves sampling of infected individuals, requiring extensive specimen acquisition, clinical testing, and sequencing coordinated across different sites and laboratories. This type of clinical surveillance is expensive, time-consuming, and subject to bias owing to disparities in public participation and frequency of testing and sequencing, which may limit outbreak preparedness and response by public health organizations, especially in underserved communities (*7*). Although clinical surveillance will remain fundamental to infectious disease response, wastewater-based approaches enable fast and cost-effective surveillance, even in current blind spots.

Wastewater monitoring enables rapid pathogen detection and community prevalence quantification. At sites where wastewater from the population collects and mixes, so too do a diverse array of microbes shed from individuals. The resulting mixture is representative of local infections, and with now-standard laboratory processing and quantitative polymerase chain reaction (qPCR), sensitive concentration measurements of specific pathogens can be obtained affordably and in only a few hours. Pathogen concentrations accurately estimate prevalence (the number of current infections in the population), and given that wastewater trends often precede corresponding clinical detections, they may allow for early detection (*8*).

Wastewater can be used to track infectious disease dynamics from the community level to building level, and from sources ranging from sewers and wastewater treatment plants (WWTPs) to surface waters and point sources (e.g., natural pooling sites) (*9*). At large WWTPs and other sites of converging wastewater flows, a single sample can capture the community-wide pathogen landscape, even for cities with millions of inhabitants (*10*). Such approaches are cost-effective and time-efficient and maximize the number of individuals covered by surveillance. In many countries, WWTPs already collect composite wastewater samples (taken regularly throughout the day) that can be analyzed to yield essential information on local pathogen prevalence, which enables timely mobilization of public health interventions. In areas lacking centralized sewer infrastructure, similar methods can be used to study samples from surface waters and point sources, although additional considerations of topography, water pooling, and flow are needed to maximize catchment size (*11*).

At a focused spatial scale, wastewater can be used for monitoring at the level of individual or small clusters of buildings to enable reliable detection of even a single infected person (*10*, *12*). Such monitoring is of particular importance to vulnerable communities and high–population density sites, such as health care (e.g., nursing homes) or educational facilities (*12*), as well as airports, where detection can be acted on to contain pathogen spread.

Understanding the determinants of pathogen incidence, however, including mutational changes, variant introductions, or emerging pathogens, requires genomic sequencing. Unlike qPCR, sequencing is effective regardless of pathogen variant and identifies mutation frequencies across the entire genome, even for mutations that exist at low levels in the population. During the COVID-19 pandemic, wastewater sequencing–based analyses have enabled early detection of emerging SARS-CoV-2 variants, estimation of variant prevalence (*10*, *13*), identification of the impact of specific mutations on pathogen fitness, and characterization of the mutational processes that lead to variants of concern

[1]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA. [2]Department of Pediatrics, University of California San Diego, La Jolla, CA, USA. [3]Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA. [4]Department of Bioengineering, University of California San Diego, La Jolla, CA, USA. Email: jolevy@scripps.edu; andersen@scripps.edu

PHOTO: ERIK JEPSEN/UC SANDIEGO

(*14*). Sequencing of building-level wastewater samples has enabled extraction of single genomes that have been used to identify new mutations and variants; reveal cryptic, or undetected, spread (especially at sites that are not well covered by clinical surveillance); and elucidate local transmission networks (*10*). The ease of sharing sequencing data enables collaborative analyses of pathogen trends around the world, enhancing preparedness and informing public health guidance.

For sequencing of a specific pathogen, targeted amplicon-based approaches are needed to selectively amplify and sequence pathogens of interest from wastewater. Although this approach presents technical challenges and requires ongoing development, amplicon sequencing can provide similar sensitivity to qPCR, excellent sequencing coverage and depth, and effectiveness regardless of pathogen variant. As a result, almost all clinical and wastewater sequencing of SARS-CoV-2 has used this targeted approach. Initiating the collection of high-quality pathogen-specific sequencing data requires only minimal training, aided by community efforts such as ARTIC that provide open-source amplicon scheme designs and protocols.

Although targeted approaches are excellent for monitoring known pathogens of concern, they examine only a tiny fraction of microbes in wastewater. For a broad survey of pathogens, untargeted metagenomic and metatranscriptomic shotgun sequencing approaches can be used to identify any circulating microbial DNA or RNA. A single sample can indicate the presence of viruses including monkeypox and influenza, identify pathogenic and antibiotic-resistant strains of bacteria (*15*), detect protist parasites including *Plasmodium falciparum* (which causes malaria), and search for new pathogens, although standard approaches may not provide sufficient sensitivity to detect rare pathogens. Unlike amplicon-based surveillance, which can identify pathogen variants shed by a handful of individuals within a population of millions (*10*, *14*), untargeted methods require greater prevalence for reliable detection, although increasing sequencing depth can help improve the likelihood of pathogen detection. For more detailed profiling of a specific class of pathogens (e.g., viruses), probe-based enrichment methods can be used prior to shotgun sequencing for "semi-targeted" study of the group of interest and may be necessary to detect rare pathogens.

Designing an effective wastewater sequencing strategy gives rise to a fundamental trade-off between speed, cost-effectiveness, and local deployability. To keep costs down, most current sequencing is performed with hundreds to thousands of samples in parallel, requiring expensive machinery, automation infrastructure (liquid-handling robots), and labor-intensive processing, with ~2- to 4-week turnaround times. However, because wastewater collections capture community-wide trends, samples can be prioritized for low-throughput, fast-turnaround sequencing at higher cost per sample for real-time tracking of pathogen dynamics. For example, highly automated fast-turnaround sequencers enable automation of the entire process from library preparation to analysis, require minimal technical expertise, and permit 24-hour sample receipt to data turnaround, but lead to a 10-fold increase in per-sample cost.

For deployment in low-resource settings, sequencing workflows need to be reconfigured in accordance with available instruments, reagents, funding, and facilities. Portable low-throughput sequencers can provide results much faster, are low cost, do not require automation infrastructure, and can be deployed almost anywhere. However, they generally have higher sequencing error rates and require additional laboratory and bioinformatics training. In low- and middle-income countries, where reagents generally need to be imported from abroad, transport is often slow and expensive, and reagents requiring refrigeration can degrade during shipping or at facilities with unstable electricity supply. Progressing toward more equitable and sustainable surveillance will require continued development of local, self-sustaining scientific ecosystems through laboratory and computational methods development and training, capacity building efforts, and financial support of domestic scientific enterprise.

Ongoing method development is key to further expanding the capabilities of wastewater surveillance. At present, untargeted metagenomic and metatranscriptomic methods suffer from bias toward a limited number of dominant bacterial species in wastewater and will require new laboratory techniques to ensure reliable detection of pathogens at low concentration. The latest enrichment methods for semi-targeted sequencing of pathogens from wastewater are specific to viruses and will need to be redesigned and tested for other types of pathogens. Current computational methods are still unable to identify new variants from community-wide mixtures and are just beginning to produce a wider range of informative readouts, including growth rates, which are essential to evaluate the potential size and speed of an emerging outbreak. As these methods evolve, there will also be a need for new standardized protocols and computational workflows that are open source and easily reproducible anywhere in the world.

The COVID-19 pandemic has laid bare the importance of pathogen surveillance to public health, and wastewater as one of our most powerful assets. National and regional wastewater surveillance programs are rapidly being established across the globe, allowing for continued analysis of SARS-CoV-2 variant dynamics despite limited clinical monitoring. Expansion of multipathogen and metagenomic wastewater sequencing efforts is enabling broad pathogen detection and genomic characterization, including environmental contamination by *Vibrio cholerae* (which causes cholera) in sites without effective sewage treatment, human and livestock-borne rotavirus (which causes gastroenteritis in children), and vaccine-derived polioviruses recently observed in London and New York. Beyond infectious pathogens, wastewater also has the potential to reveal changes in the human gut microbiome, which have been correlated with a wide range of health conditions and disease risks.

There is now an opportunity to build upon existing momentum to form the backbone of future surveillance capacity and scientific ecosystems across the world. Doing so will require global surveillance networks to encourage equitable technology distribution, data sharing, and collective exploration of microbial diversity, both human and zoonotic, to help pinpoint possible outbreaks and spillover risks. Close integration with local and international public health organizations will be essential to ensure timely, transparent, and effective intervention. Wastewater provides the means to identify, prepare for, and respond to future pandemic threats—what remains is to expand capacity to ensure readiness for the next potential outbreak. ■

REFERENCES AND NOTES

1. A. E. Kirby *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 1242 (2021).
2. F. Amman *et al.*, *Nat. Biotechnol.* 10.1038/s41587-022-01387-y (2022).
3. M. K. Wolfe *et al.*, *Environ. Sci. Technol. Lett.* **9**, 687 (2022).
4. S. Wurtzer *et al.*, bioRxiv 10.1101/2022.08.18.22278938 (2022).
5. L. Ma *et al.*, *Environ. Sci. Technol.* **50**, 420 (2016).
6. S. R. Finkbeiner *et al.*, *PLOS Pathog.* **4**, e1000011 (2008).
7. A. F. Brito *et al.*, medRxiv 10.1101/2021.08.21.21262393 (2021).
8. S. Karthikeyan *et al.*, *mSystems* **6**, e00045-21 (2021).
9. S. Lamba *et al.*, medRxiv 10.1101/2022.07.14.22277616 (2022).
10. S. Karthikeyan *et al.*, *Nature* **609**, 101 (2022).
11. C. B. Uzzell *et al.*, medRxiv 10.1101/2021.05.21.21257547 (2021).
12. R. K. Fielding-Miller *et al.*, medRxiv 10.1101/2021.10.19.21265226 (2021) .
13. K. Jahn *et al.*, *Nat. Microbiol.* **7**, 1151 (2022).
14. D. A. Gregory *et al.*, medRxiv 10.1101/2022.06.03.22275961 (2022).
15. A. Harrington *et al.*, *Sci. Total Environ.* **853**, 158577 (2022).

ACKNOWLEDGMENTS

The authors are supported by National Institutes of Health (NIH) National Institute of Allergy and Infectious Diseases U01AI151812 (K.G.A), National Center for Advancing Translational Sciences UL1TR002550 (K.G.A), Centers for Disease Control and Prevention contracts 75D301-20-C-09795 (K.G.A., R.K.) and 75D301-22-R-72097 (K.G.A, R.K.), and NIH 5T32AI007244-38 (J.I.L.).

10.1126/science.ade2503

AGING

Obesity confers macrophage memory

Epigenetic programming of myeloid cells in obesity contributes to macular degeneration

By **Kevin D. Mangum** and **Katherine A. Gallagher**

Obesity is an important and often modifiable risk factor that induces systemic inflammation through reprogramming of hematopoietic cells. Evidence suggests that environmental factors such as free fatty acids, which are present in high amounts in obesity, train innate immune cells such as monocytes and macrophages to an inflammatory phenotype (*1*–*3*). However, the mechanisms that alter immune cells in this setting are unclear. On page 45 of this issue, Hata *et al.* (*4*) report that free fatty acids associated with obesity epigenetically alter adipose-resident macrophages (ARMs) toward a proinflammatory phenotype, which is retained through "epigenetic memory" during the aging process in mice. These patterned macrophages then travel to the eye, where they initiate an inflammatory program that promotes age-related macular degeneration (AMD).

Hata *et al.* found that ARMs from mice fed a high-fat diet exhibit increased expression of inflammatory genes, including tumor necrosis factor (*Tnf*), interleukin-1β (*Il1b*), and *Il6*, which persists after mice return to normal weight. The authors show that this lasting proinflammatory state is mediated by epigenetic changes that occurred in ARMs during the obese state. These epigenetic alterations facilitate open chromatin at binding sites for the activator protein-1 (AP-1) family of transcription factors, which allows expression of inflammatory genes, thereby promoting pathologic inflammation (see the figure). Using a laser-induced model of AMD, the authors find that previously obese mice exhibited increased choroidal neovascularization, a feature of AMD, compared with mice that were never obese and maintained a normal weight throughout their life span. Hata *et al.* demonstrate that depletion of both local and distal myeloid populations in mice that were once obese normalizes retinal dysfunction and limits the development of AMD, indicating that both local retinal and systemic ARMs contribute to AMD.

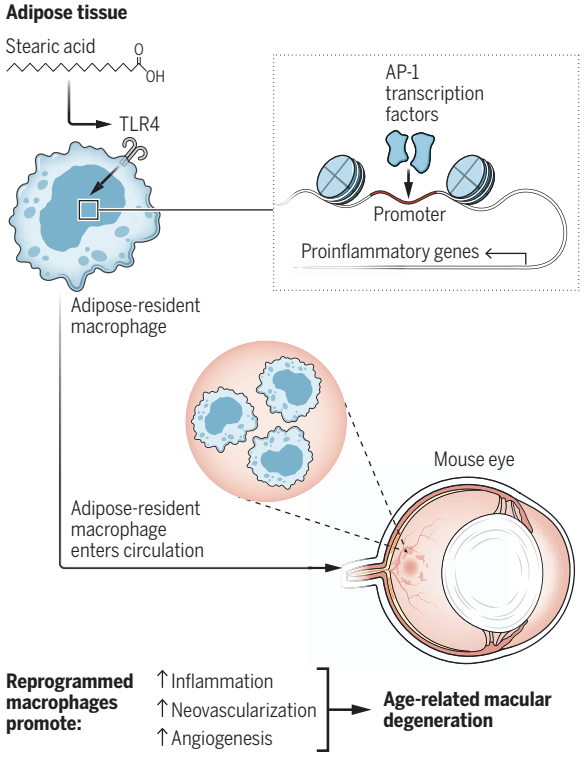
The concept of epigenetic memory, in which cells stably retain the ability to transmit specific gene expression patterns to cell progeny, has been increasingly observed in several disease states, such as cancer and cardiovascular disease (*5*–*8*). The exact molecular mechanisms that result in epigenetic memory are unclear. Studies have linked myeloid cell–mediated inflammation to the Toll-like receptor (TLR) family, which "senses" danger signals, or, in this case, fatty acids, and transduces them into downstream cellular signaling cascades that induce inflammatory gene expression and a metabolic shift that poises immune cells toward an inflammatory response to further stimulus. Oxidized low-density lipoprotein (LDL) particles, which occur at high concentrations in the circulation of people with obesity, induce an innate immune response and confer epigenetic memory in monocytes and myeloid precursors in bone marrow and spleen that influences inflammation and disease (*1*). In this manner, myeloid cells from mice that are fed a high-fat diet, even when switched to a normal diet, exhibit augmented TLR responses.

In a metabolic disease state, such as diabetes, myeloid cells are epigenetically reprogrammed distally in bone marrow and the spleen, and then, through chemokines released in the local tissue environment, these cells traffic to their respective sites (e.g., wounds, kidneys, heart) with an altered phenotype, where they regulate disease (*6*, *9*, *10*). In the example of cutaneous tissue injury, peripheral monocytes are recruited to the wound, where they first differentiate into proinflammatory macrophages—which secrete inflammatory cytokines such as IL-1β, TNF, and IL-6—and then transition to an anti-inflammatory phenotype that promotes tissue repair (*11*). However, in diabetes, macrophages are epigenetically altered at distal sites, and once recruited to the tissue, they exhibit a sustained inflammatory response that interferes with wound repair and prevents them from transitioning to an anti-inflammatory macrophage phenotype (*12*). Consistently, Hata *et al.* show that resident inflammatory cells in adipose tissue can be epigenetically reprogrammed and later traffic to distant sites (the eye) and locally regulate disease severity.

The study of Hata *et al.* raises important questions about the upstream pathways that are responsible for epigenetic reprogramming in macrophages and whether targeting these pathways can reverse epigenetic changes. Although the

Epigenetic reprogramming of macrophages

Fatty acids, such as stearic acid, present in the adipose tissue of obese mice confer epigenetic alterations in adipose-resident macrophages through Toll-like receptor 4 (TLR4)–mediated chromatin opening that recruits activator protein-1 (AP-1) transcription factors. These epigenetically reprogrammed macrophages traffic to the eye, where they promote pathological inflammation and age-related macular degeneration, even after mice return to normal weight.



GRAPHIC: A. FISHER/*SCIENCE*

Departments of Surgery and Microbiology and Immunology, University of Michigan, Ann Arbor, MI, USA. Email: kgallag@med.umich.edu

authors selectively depleted ARMs in mice to inhibit their functional effect on AMD, the therapeutic potential of this strategy is limited by off-target effects, because macrophages in adipose tissue are likely important for maintaining adipose tissue homeostasis and performing compensatory functions elsewhere. Accordingly, as intricacies of these epigenetic-based mechanisms are further uncovered, it will be necessary to address how to design effective, cell-specific treatments. Likely strategies involve cell-specific targeting of epigenetic enzymes themselves, interactions between transcription factors (e.g., AP-1 family) and epigenetic enzymes (e.g., histone acetyltransferases), or upstream signaling cascades that regulate epigenetic alterations. Other possible targets include the transcriptional regulators identified by studies of epigenetic patterning induced in myeloid cells by a high-fat diet, such as the signal transducers and activators of transcription (STAT) family, nuclear factor-κB (NF-κB), interferon regulatory factor (IRF), and JUN (*1*).

Elucidating the specific upstream mechanism(s) that promote the open chromatin conformation also warrants further investigation. Moreover, interrogation of TLR signaling may help uncover how alterations in this pathway affect cellular and metabolic shifts in ARMs that are responsible for their dysregulated function in disease. For example, does activation of TLR signaling alone confer epigenetic memory in macrophages? And if so, what is the overlap between TLR-induced regions of open chromatin and those found in ARMs from obese mice? According to the study of Hata *et al.*, substantial overlap between these regions might be expected; however, nonoverlapping regions might be used to identify alternate upstream regulators of myeloid patterning. Additionally, the specific contribution of tissue-resident macrophages versus bone marrow–derived myeloid cells to inflammatory diseases should be dissected further to guide more effective and targeted therapeutics. ■

REFERENCES AND NOTES

1. A. Christ *et al.*, *Cell* **172**, 162 (2018).
2. K. Ganeshan, A. Chawla, *Annu. Rev. Immunol.* **32**, 609 (2014).
3. M. G. Netea *et al.*, *Science* **352**, aaf1098 (2016).
4. M. Hata *et al.*, *Science* **379**, 45 (2023).
5. L. J. Gaydos, W. Wang, S. Strome, *Science* **345**, 1515 (2014).
6. G. J. Koelwyn *et al.*, *Nat. Med.* **26**, 1452 (2020).
7. A. S. Kimball *et al.*, *Immunity* **51**, 258 (2019).
8. F. M. Davis *et al.*, *J. Exp. Med.* **218**, e20201839 (2021).
9. M. Tomczyk *et al.*, *Basic Res. Cardiol.* **112**, 39 (2017).
10. T. Tran *et al.*, *Nat. Commun.* **14**, 13 (2022).
11. T. A. Wynn, K. M. Vannella, *Immunity* **44**, 450 (2016).
12. J. Yan *et al.*, *Nat. Commun.* **9**, 33 (2018).

10.1126/science.adf6582

CLIMATE

Acting now will reduce glacier loss

Many of the world's glaciers will disappear, but quick action will make a difference

By **Guðfinna Aðalgeirsdóttir**[1] *and* **Timothy D. James**[2]

As global mean temperature rises in pace with increasing greenhouse gas emissions, the future of the world's glaciers looks bleak. Rates of glacier mass loss have increased over the past two decades (*1*), a trend that will continue even if emissions are capped (*2*–*4*). Despite their small size relative to the Greenland and Antarctic ice sheets, these dwindling ice stores are important. They currently contribute as much to sea level as the ice sheets (*3*), their disappearance means water insecurity for millions, and their retreat increases glacier hazard frequency, such as glacier outburst floods and landslides (*2*). Although most countries have agreed to pursue temperature limits within 1.5°C above preindustrial levels (*5*), these targets are unmet (*6*). On page 78 of this issue, Rounce *et al.* (*7*) present a model of the fate of all 215,547 glaciers under different climate scenarios. Their findings emphasize the need to act now to prevent substantial glacier loss.

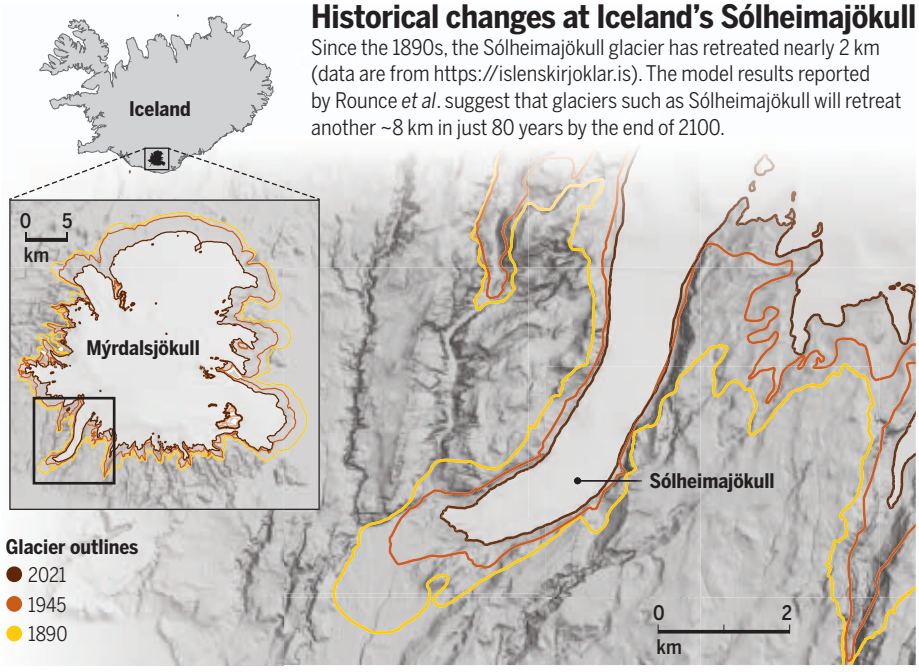
The path to a global projection model for individual glaciers has been long. Historically, glacier measurements were made directly, on the surface, and later from aircraft using aerial photography. Both methods are labor intensive and costly, and, therefore, they can only be undertaken consistently and over long time periods for a handful of accessible glaciers (*8*). Nonetheless, these records provided important early evidence of the sensitivity of glaciers to changes in climate and of their dynamic nature. Since then, satellite data provide records of the number and area of glaciers, elevation changes, their velocity, and measurements of mass changes for essentially every glacier on Earth. Although these records are insufficient for understanding either the effects of climate changes from the more distant past or how glaciers will change in the future, they have provided key calibration data for global glacier models that can help provide a better understanding of both.

Global models projecting glacier response to climate are fairly recent, evolving from relatively simple volume-change projections; to more complex models with different approaches for representing melt, accumulation, frontal ablation, and geometric changes; to the ambitious ensemble projects, such as the Glacier Model Intercomparison Project (Glacier-MIP) (*4*, *9*). Most models have represented glacier dynamics—the important component of glacier mass balance involving ice flow and iceberg calving—using basic assumptions about mass redistribution or the relationship between glacier volume and area. Until now, only one of these global models (*10*) has simulated glacier dynamics using a flowline model, which is a more realistic way to model glacier changes, where each glacier is represented by a centerline along its length. Despite recent unprecedented efforts in global glacier projections, several key physical processes that control glacier mass loss remain absent.

The study by Rounce *et al.* is a substantial step forward in predictive modeling. They provide a set of predictions for every glacier on Earth, covering 2015 to 2100. These are based on reference future greenhouse gas emission (representative concentration pathways) (*11*) and socioeconomic (shared socioeconomic pathways) (*12*) scenarios, taking advantage of new datasets of global glacier mass balance (*1*) and near-global frontal ablation (*13*). Their model represents glacier dynamics with a flowline model and includes important physical processes, such as the effect of surface debris on melt and frontal ablation. Although there is room for improvement, with more-accurate glacier thickness estimates and higher-resolution climate condition forecasts, this model is the most comprehensive so far.

By providing model results in the context of policy-relevant end-of-century mean global temperature increases, the authors directly attribute regional mass loss, sea level contributions, and the number of lost glaciers to the consequences of meeting and failing to meet the Paris Agreement's 1.5° to 2°C temperature limit, and they tell a tragic tale. Even in the optimistic best-case

[1]Institute of Earth Sciences, University of Iceland, Reykjavík, Iceland. [2]Department of Geography and Planning, Queen's University, Kingston, ONT, Canada. Email: gua@hi.is; tdj1@queensu.ca


Historical changes at Iceland's Sólheimajökull
Since the 1890s, the Sólheimajökull glacier has retreated nearly 2 km (data are from https://islenskirjoklar.is). The model results reported by Rounce et al. suggest that glaciers such as Sólheimajökull will retreat another ~8 km in just 80 years by the end of 2100.
Iceland
0 5
km
Mýrdalsjökull
Sólheimajökull
Glacier outlines
2021
1945
1890
0 2
km

scenario, where global mean temperature is kept within 1.5°C of preindustrial levels, the results of Rounce *et al.* suggest that nearly 60% of the world's glaciers will be lost by 2100 (see the figure). Within 2°C, the agreed warming limit of the Paris Agreement, ~70% of glaciers <1 km^2 and nearly 20% of glaciers between 1 and 10 km^2 will be lost. However, estimates from COP26 predict that global mean temperature increase is headed toward ~2.7°C by 2100 (*6*). At these temperatures and higher, rates of mass loss will increase even for the large glaciers in the polar regions that remain largely insensitive to increases below 2°C. Importantly, even if temperature is stabilized at any level, the glaciers that remain will continue losing mass beyond the model's 2100 end point. Taken together with forecasts predicting the loss of summer Arctic sea ice, increasing melt from Greenland and Antarctica, rising sea levels (*3*), and an alarming decline of biodiversity (*14*), the "climate emergency" has never been clearer.

Despite this urgency, national climate action plans to cut emissions and adapt to climate impacts (submitted ahead of COP26) continue to be insufficient to achieve the temperature goal of the Paris Agreement (*6*). Disappointingly, substantial attempts were made at COP27 to weaken ambitions to quickly cut greenhouse gas emissions, and the negotiations ended without a commitment to phase out fossil fuels. We seem to be stuck. Although the causes of these failures are complex, maybe scientists could help by doing a better job with messaging. Recent studies have revealed that positive opportunistic framing of climate change is more effective at inciting public support for climate policies than focusing on urgency and consequences (*15*). Rounce *et al.*, while issuing a stark warning about the consequences of insufficient action, achieve this framing with an important message: Although it is too late to avoid losing many glaciers, any effort to limit global mean temperature rise will have a direct effect on reducing how many glaciers will be lost. Although projections provide important insight into the magnitude, patterns, and timing of changes that are key for adaptation and mitigation, when presented in a way that translates any climate action into observable outcomes, these projections can provide a key motivational message that is needed in this critical decade for climate action. ■

REFERENCES AND NOTES

1. R. Hugonnet *et al.*, *Nature* **592**, 726 (2021).
2. The Intergovernmental Panel on Climate Change (IPCC), "IPCC Special Report on the Ocean and Cryosphere in a Changing Climate," H.-O. Pörtner *et al.*, Eds. (Cambridge Univ. Press, 2019).
3. IPCC, "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change" (Cambridge Univ. Press, 2021).
4. B. Marzeion *et al.*, *Earth's Future* **8**, e2019EF001470 (2020).
5. United Nations (UN)/Framework Convention on Climate Change, "Adoption of the Paris Agreement, 21st Conference of the Parties" (UN, 2015).
6. UN Environment Programme (UNEP), "Emissions Gap Report 2021: The Heat Is On – A World of Climate Promises Not Yet Delivered" (UN, 2021).
7. D. R. Rounce *et al.*, *Science* **379**, 78 (2023).
8. M. Zemp *et al.*, *Nature* **568**, 382 (2019).
9. R. Hock *et al.*, *J. Glaciol.* **65**, 453 (2019).
10. F. Maussion *et al.*, *Geosci. Model Dev.* **12**, 909 (2019).
11. D. P. van Vuuren *et al.*, *Clim. Change* **109**, 5 (2011).
12. B. C. O'Neill *et al.*, *Clim. Change* **122**, 387 (2014).
13. W. Kochtitzky *et al.*, *Nat. Commun.* **13**, 5835 (2022).
14. World Wildlife Fund (WWF), "Living Planet Report 2022: Building a Nature-Positive Society" (WWF, 2022).
15. N. Dasandi *et al.*, *Commun. Earth Environ.* **3**, 239 (2022).

10.1126/science.ade2355

OPTICS

Toward ultrathin optics

Function determines the minimum thickness of an optical system

By **Francesco Monticone**

Optical systems are ubiquitous in today's world, from smartphone cameras to microscopes and telescopes, all capturing ever-increasing amounts of data and augmenting our visual perception. Making lenses and other optical components thinner has been an important goal in the field of optics in recent years. Creating thin and lightweight virtual and augmented reality headsets or night-vision goggles, for example, would be a game changer for consumer and military applications. The push toward miniaturization requires better theoretical understanding of how thin an optical system can be. On page 41 of this issue, Miller (*1*) derives quantitative limits on the minimum possible thickness of an optical system, based only on the function that the optics is to perform. This finding and other recent results (*2*–*9*) address foundational questions regarding why optical systems require certain thicknesses.

Any optical system, such as a virtual reality headset or a smartphone camera, is composed of a combination of optical elements (usually lenses), optoelectronic devices (such as an imaging sensor or a display), and empty spaces between them. In a simple imaging system, such as the human eye or a camera, a lens bends light rays, which then converge in the space behind the lens, focusing onto the electronic sensor, film, or retina where the image is formed. The optical behavior of such a system, and its physical limits, may be better described by considering the wave nature of light. The physics of waves explains the natural tendency of light to spread out, especially around obstacles and through apertures—a phenomenon known as diffraction, which was first systematically observed and documented by the Italian physicist Francesco Maria Grimaldi in the 17th century. For a focusing lens, the wavefront emerging from

School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA.
Email: francesco.monticone@cornell.edu

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

the lens, which may be interpreted as the superposition of secondary spherical waves, converges at a certain distance where a spot of concentrated light is formed. Because of diffraction through the finite lens aperture, the width of this focal spot, which determines the resolution of the imaging system, depends on the area of the lens and the finite wavelength of light. The wave nature of light therefore explains the fundamental need for area in optics to achieve a certain imaging resolution.

The reason why optics may also require a nonzero thickness is more involved. This can be explained by first recognizing that, in most optical systems, such as a camera, the output of the system at a certain point (e.g., the light intensity on a sensor pixel) depends on the input not just at a single point but over an extended region (e.g., the light on the entire lens surface). This type of input-output relationship is known as nonlocal and implies that the optical system may need to route waves sideways from a portion of the input surface to a point on the output surface, crossing a transverse aperture (see the figure). Mathematically, this can be interpreted as a communication problem between the two surfaces, where a certain number of channels are needed for such communication. Following these considerations, Miller showed that solely on the basis of the function that the optical system is to perform (e.g., imaging), one can calculate the number of independent channels that must cross the transverse aperture and then determine the size of the aperture needed to accommodate the number of channels, assuming that the system only contains transparent, nonabsorbing materials. This last step is related to light diffraction because the natural spreading of waves makes it difficult to confine light in channels with a small cross section relative to the wavelength. This implies that the size of the transverse aperture, and therefore the thickness of the optical system, must grow in proportion to the number of channels it needs to accommodate and to the finite wavelength of light. Thus, wave physics explains not only the fundamental need for a minimum area but also the need for a minimum thickness, determined by the function of the optical system and not by its specific implementation details.

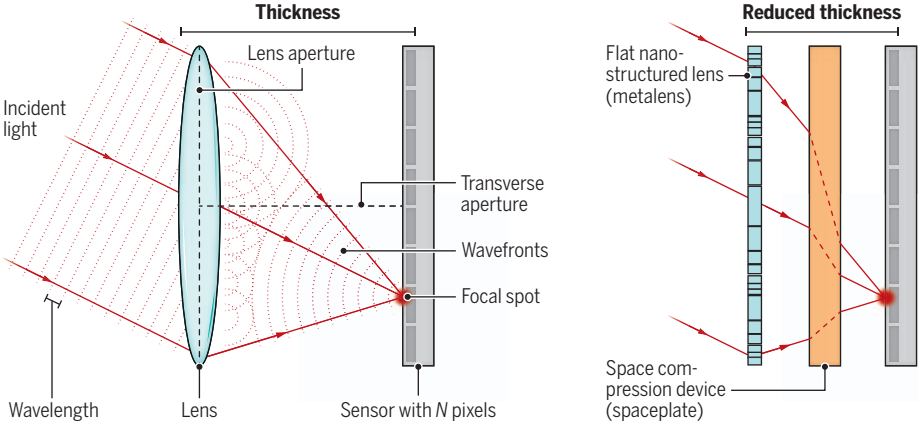
Interest in the ultimate thickness limits of optics is motivated by recent advances in the quest to miniaturize optical systems. Engineered nanostructured films known as metasurfaces or metalenses, for instance, have the potential to replace conventional refractive lenses and enable much greater design flexibility (*10*). However, even after replacing all optical elements with ultrathin metasurfaces, a typical optical system would still require large empty spaces to let light propagate—converging or diverging according to the desired functionality. Miniaturization of these spaces has been a focus within the emerging field of nonlocal flat optics, which seeks to enhance the nonlocal response inherent in wave-based systems (*11*, *12*). Nonlocal effects are engineered to create planar optical systems that are substantially thinner or are endowed with new functionalities. Notably, these ideas are the basis of new optical elements known as spaceplates (*8*, *11*, *13*–*15*), which are designed to mimic the way that light propagates and spreads in empty space but over a shorter distance. With a spaceplate, light effectively experiences more space than what actually exists. The combination of metalenses (replacing standard lenses) and spaceplates (compressing space for light propagation) may lead to substantial miniaturization of any optical system. This could enable, for example, a paper-thin camera with large optical magnification despite its small thickness. In these future systems, the optics may be implemented as a planar ultrathin structure integrated directly on top of an electronic sensor.

Miller's study provides a general approach to understanding how thin these systems can ultimately become. For example, Miller directly applied his analysis to spaceplates, showing that the derived limits are obeyed and approached by existing designs (*15*). The same approach was also used to determine the minimum required thickness for imaging systems as well as for metasurfaces that perform analog computations and image processing operations (*11*). Based on different methods, other recent studies have addressed the question of minimum required thickness for other classes of devices that fall beyond the purview of Miller's study, including light absorbers (*3*, *4*), metasurface reflectors (*4*, *9*), and full-color metalenses (*6*, *7*).

As the race toward creating ultrathin optical systems is expected to accelerate in the coming years, new concepts and cutting-edge technologies, such as metasurfaces and spaceplates, and new approaches to understanding their fundamental limits will play increasingly important roles. These advances may pave the way for a new generation of optical systems, finally approaching—after nearly four centuries since the first systematic observations of light diffraction—the ultimate thickness limits allowed by wave physics. ■

How thin can an optical system get?

In a lens-sensor imaging system (left), a certain area and a certain distance between the focusing lens and the sensor are needed. A combination of new optical elements (right) allows a reduction in thickness. This may lead to the miniaturization of a wide range of optical systems, approaching their ultimate thickness limits.



GRAPHIC: A. MASTIN/*SCIENCE*

REFERENCES AND NOTES

1. D. A. B. Miller, *Science* **379**, 41 (2023).
2. D. A. B. Miller, *J. Opt. Soc. Am. B* **24**, A1 (2007).
3. Z. Kuang, L. Zhang, O. D. Miller, *Optica* **7**, 1746 (2020).
4. M. Gustafsson, K. Schab, L. Jelinek, M. Capek, *New J. Phys.* **22**, 073013 (2020).
5. P. Chao, B. Strekha, R. Kuate Defo, S. Molesky, A. W. Rodriguez, *Nat. Rev. Phys.* **4**, 543 (2022).
6. F. Presutti, F. Monticone, *Optica* **7**, 624 (2020).
7. J. Engelberg, U. Levy, *Optica* **8**, 834 (2021).
8. K. Shastri, O. Reshef, R. W. Boyd, J. S. Lundeen, F. Monticone, *Optica* **9**, 738 (2022).
9. M. I. Abdelrahman, F. Monticone, *Adv. Optical Mater.* 10.1002/adom.202201782 (2022).
10. P. Lalanne, P. Chavel, *Laser Photonics Rev.* **11**, 1600295 (2017).
11. K. Shastri, F. Monticone, *Nat. Photon.* 10.1038/s41566-022-01098-5 (2022).
12. A. Overvig, A. Alù, *Laser Photonics Rev.* **16**, 2100633 (2022).
13. C. Guo, H. Wang, S. Fan, *Optica* **7**, 1133 (2020).
14. O. Reshef *et al.*, *Nat. Commun.* **12**, 3512 (2021).
15. A. Chen, F. Monticone, *ACS Photonics* **8**, 1439 (2021).

10.1126/science.adf2197

POLICY FORUM

Soybean is harvested in Brazil for animal feed in Europe and China, where nitrogen and other nutrients accumulate. This is an example of worldwide connections of soils and soil health.

PHOTO: SILVIO AVILA/AFP/GETTY IMAGES

SOIL HEALTH

Soil biodiversity needs policy without borders

Soil health laws should account for global soil connections

By **Wim H. van der Putten[1,2], Richard D. Bardgett[3], Monica Farfan[4,5], Luca Montanarella[6], Johan Six[7], Diana H. Wall[4,5]**

Soil biodiversity is crucial for healthy soils, on which we all depend for food, human health, aboveground biodiversity, and climate control. It is well known that land use intensification, climate change, environmental pollution, and mining activities degrade soil biodiversity. However, most current and intended policies on soil protection not only lack a holistic view on how biological, physical, and chemical components of soil health are integrated but also overlook how soils across national borders and continents are connected by human activities. The challenge is to use recent advancements in understanding the distribution and functional roles of soil biodiversity in developing policy on restoring and protecting soil health across borders. Thus, policy should focus not only on soils within a nation or union of nations but also on preventing negative footprints on each other's soils.

Numerous factors—such as urbanization, automation, disease outbreaks, natural disasters, and even wars—influence how land is used, which affects the capacity of soils to perform multiple functions, also called soil health (*1*). Searching for sustainable land use while providing food and feed for a more demanding population and dealing with growing demands on land for multiple other functions requires insights into the many factors that influence land use. Often, land use options are considered trade-offs, and the challenge is to search for win-win options, for example, climate change mitigation by biodiversity restoration. A transdisciplinary approach may help to understand possibilities and trade-offs to achieve a more sustainable society (*2*). Although an awareness that healthy soils are the basis of a healthy society is growing, anchoring this view into policy is still a challenge.

Soil protection requires an integrated legal framework to address the multitude of processes that are involved in land degradation, but most existing soil laws that should protect soils now focus on single issues, such as desertification or soil contamination. Moreover, soil protection laws are mostly national (*3*), although soil protection does not stop at national frontiers. For instance, current climate change caused by poor land use and industrialization outside sub-Arctic regions causes melting of the permafrost, which in turn exacerbates climate change through the release of carbon dioxide and methane to the atmosphere.

Although it is widely acknowledged that plants, birds, butterflies, and many other animal species need to be protected, little explicit attention exists for protecting soil biodiversity (*4*). The European Union (EU) Soil Strategy for 2030 (*5*) has been set up to combat declining soil health in Europe and beyond. The ambition is to have healthy soils in the entirety of Europe by 2050. An important aspect of this ambition is that the EU is planning to propose a binding European Soil Health Law in 2023. To the best of our knowledge, this is the first and most inclusive soil health protection law that recognizes the ecosystem services provided by healthy soils and the need to protect those services for future generations. Proposing a soil health law is an important step toward a sustainable society; however, the real challenge is to make it work.

To make the EU's Soil Health Law operational, soil health needs to be measurable. Different from soil quality, which is largely chemical in focus and mostly used to characterize the status of soil to sustain crop productivity, soil health is a more holistic concept (*6*). It is based on the recognition of the ecosystem services that soils provide. As defined in the EU soil strategy, soils are healthy when they are in good chemical, biological, and physical condition and are able to continuously provide as many of the ecosystem services as possible. Soil health addresses the sustainability goals set by the United Nations (UN), which have been adopted by many countries. However, finding effective, easy-to-measure indicators for soil health is challenging, because there is no one-size-fits-all indicator for all circumstances, just as in the case of soil quality (*7*).

STRUGGLE FOR INDICATORS

Measuring soil health requires information on biological, chemical, and physical properties of soil, the obtainment of which is a substantial effort that will be too costly for

[1]Netherlands Institute of Ecology, Wageningen, Netherlands. [2]Department of Nematology, Wageningen University, Wageningen, Netherlands. [3]Department of Earth and Environmental Sciences, The University of Manchester, Manchester, UK. [4]School of Global Environmental Sustainability, Colorado State University, Ft. Collins, CO, USA. [5]Department of Biology, Colorado State University, Ft. Collins, CO, USA. [6]European Commission, Joint Research Centre, Ispra, VA, Italy. [7]Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland. Email: w.vanderputten@nioo.knaw.nl

individual landowners. For example, soils contain an immense amount of biodiversity. One handful of soil might contain more than 5000 taxa, including species of viruses, bacteria, archaea, fungi, protists, nematodes, earthworms, and other small invertebrates. These organisms feed on live plant material, organic debris, and on each other as part of an intricate belowground food web. Analyzing the full soil biodiversity of every piece of land at time intervals that are realistic for soil biota is costly. Therefore, the EU Soil Strategy has proposed an EU-wide scheme to enable landowners to perform soil analysis with the "test your soil for free" initiative.

These free tests are a great preliminary step, but the numerous tests that are available are not all equally well calibrated. Because soil biota perform many different functions, there is no species that can represent all. Some species provide soil structure by converting plant litter into plant-accessible nutrients and others determine whether soils produce or consume greenhouse gases, whereas certain soil biota provide plants with protection against natural enemies, both below and above the ground (*6*). In addition, mutualistic symbionts, such as mycorrhizal fungi and nitrogen-fixing bacteria, play an important role in enhancing soil fertility, and they may protect plants against soilborne pathogens and root-feeding herbivores.

Likewise, although soil types have been relatively well described at a regional scale, local soil conditions can be highly variable. For example, soils are made up of different layers, or horizons, with contrasting physical, chemical, and biological characteristics. Proper quantification of biological, physical, and chemical soil properties at realistic spatial and temporal scales requires intensive soil sampling and laboratory analyses. Implementation of this information in a soil health law requires indicator values that are as simple as possible and can be used to propose measures to be implemented locally. For instance, the proposed establishment of EU-wide Soil Health Districts will enable a locally adapted approach. The first precursors of these districts are the 100 Living Labs and Lighthouses to be established within the research mission of the EU named "A Soil Deal for Europe." This regionalized consideration of soils mirrors a similar successful approach in the United States that established operational Soil Conservation Districts, which have been the key to the success of the US Soil Conservation Act of 1935. This act reversed the "dust-bowl syndrome" that caused displacement and hunger for millions of Americans.

Farmers are strongly focused on controlling pathogens and root feeders using chemical crop protection, crop rotation, or intercropping. However, excessive use of chemical fertilizer and soil tillage, as well as narrow crop rotation, enhance the abundance of crop growth–reducing plant-parasitic nematodes but reduce the abundance of most beneficial soil fungi (*8*). This makes intensive agriculture vulnerable to extreme drought and rainfall events, which increases the need for irrigation and drainage of soils. Meanwhile, fertilizers, biocides, and mechanical soil tillage are increasingly needed to grow high-yielding crop varieties. Implementation of a soil health law requires assessments that inform on the complexity and functioning of the entire soil food web. The question remains whether there are simple indicators for such complexity.

It will also be challenging to find natural soils that can act as a reference for healthy agricultural soils. In natural ecosystems, burrowing soil organisms such as earthworms promote soil structure and water infiltration, decomposing bacteria and fungi recycle organic and mineral nutrients by breaking down plant litter and root excretion products, and symbiotic and mutualistic microbes, as well as soilborne pathogens, may have a positive role in suppressing dominant plant species, thereby promoting the coexistence of multiple species in natural grasslands, forests, and other ecosystems. Moreover, belowground biota are indirectly involved in the control of aboveground enemies, including pathogenic bacteria and fungi and shoot-feeding insects. In densely populated industrialized parts of the world, it is likely that most, if not all, soil is affected by human activities. The challenge for soil laws will be how to develop gold standards for healthy soils.

ASSESSING SOIL BIODIVERSITY

The Convention on Biological Diversity defines biodiversity as the variation in life from genes to ecosystems and landscapes. This confronts comprehensive soil biodiversity monitoring with several challenges. At the genetic level, the species concept is less clear for microbes than for plants or animals, so that microbes are identified as operational taxonomic units (*9*). These approaches are often used for analyzing community composition, both of microbes and small invertebrate fauna that occur in soil. Detailed studies of soilborne pathogens, root parasites, and model organisms have revealed a high level of genetic diversity in individual species. Examples include the genetic population structure of soilborne pest and pathogen species, such as that of fungi that cause plant wilting. However, these analyses may be too detailed for general soil health assessments.

Molecular identification, through DNA or RNA barcoding, is widely used to characterize the diversity and community composition of microbes, protists, nematodes, and other taxonomic groups of soil organisms. The use of metagenomics and metatranscriptomics can assist in assessing soil microbial communities and their functional potential. However, challenges remain in developing uniformly standardized methods to conduct sample collection and analyses (*10*). This standardization is needed because expressions of units of species and individuals vary from numbers per weight or volume of soil to numbers per square meter(s), which makes comparisons across studies and taxonomic groups challenging (*11*). The use of environmental DNA (eDNA) might provide more uniformity and practicality for biodiversity assessments, provided that they are well calibrated (*12*).

Using new molecular identification techniques may help to determine which soil species can be found where and whether they are threatened or not. The current lack of knowledge hampers policy on soil biodiversity protection. It is unknown whether certain soil organisms should be placed on a "red list" of species that are threatened by extinction (*1*). There are several examples of highly sensitive ecosystems in which single soilborne species perform essential functions, such as the southernmost bacteria-eating nematode *Scottnema lindsayae*, which lives in the dry valleys of Antarctica and is threatened by climate warming. This nematode plays a crucial role in nutrient cycling.

There are also examples of invasive soil organisms, such as the oomycete *Phytophthora cinnamomi*, which destroys forest vegetation in Australia, and invasive earthworms that change nutrient cycles in North American forests. The EU soil strategy makes explicit reference to the need to assess the risk of alien flatworm species for their potential inclusion in the list of "invasive alien species of Union concern," in line with the EU Invasive Alien Species Regulation. Nevertheless, the concepts of rarity and invasiveness, which are well accepted for plants and aboveground animals, remain largely unused for soil biodiversity protection across the globe.

When considering the substantial efforts that will be required to design and undertake representative soil sampling, key questions arise: What sort of data are needed and for what purpose, and how can this information be collected and made available to all? It is likely not necessary to first identify and describe all species of soil biota before using the presently available information for biodiversity protection. An alter-

native to species identification is to qualify soil biota according to traits, such as nitrite oxidizers or cellulose degraders. This functional approach has also been used to develop food web models to analyze carbon and nitrogen flows through soils.

ENHANCING SOIL LITERACY

Despite the challenges of analyzing soil biodiversity, progress has been made in bringing soil biodiversity to a wider audience. For example, the publication of global maps of diversity within taxonomic groups—bacteria, (mycorrhizal) fungi, nematodes, microarthropods, and earthworms—has raised some awareness of the enormous diversity of life in soil and how it is distributed, as has the recent production of the Global Soil Biodiversity Atlas (*13*) and the first Global Soil Biodiversity Assessment (*1*). These publications provide the most up-to-date syntheses of the current status of soil biodiversity and have stimulated the initiation of the International Network on Soil Biodiversity (NETSOB), which aims to become a network of Global Soil Biodiversity Observatories (GLOSOBs).

How soil observatories will work and what will be measured is still under development. Ideally these will be sites where soil biodiversity is monitored over longer periods with standardized protocols to quantify consequences of global changes in climate, land use, invasive species, erosion, and other changes that require detailed and long-term monitoring efforts. These observatories will complement the global Soil Biodiversity Observation Network (SoilBON) and EU Land Use and Coverage Area frame Survey (LUCAS) program, which have started to collect soil biodiversity information from numerous sampling sites across Europe. These large-scale monitoring programs may also be linked to evaluations of the EU's Soil Health Law and national surveys and to citizen science activities on soils and soil biodiversity, such as the Dutch Soil Animal Days project. Public availability of data from such surveys will facilitate further use.

WORLDWIDE SOIL CONNECTIONS

Restoring soil health in one region also requires attention to its consequences in other regions of the globe. Worldwide, soils and soil biodiversity are connected through international trade, climate change, invasive exotic species, tourism, atmospheric composition, pollutants, and other environmental changes. Improving soil health in the EU may have planned or unplanned side effects elsewhere on the globe. Those side effects need to be monitored for policy improvement. For example, intensive animal farming in industrialized countries requires additional feed from other locations. The resulting flow of resources (e.g., nitrogen, phosphorus, potassium) may degrade soil biodiversity and soil health in both feed-production and feed-consumption locations. At feed-production locations, soil biodiversity may decline as a result of lowered soil organic matter content and export of nutrients, whereas in feed-consumption locations, excessive availability of manure results in nutrient enrichment of soil, shifts to bacteria-based soil food webs, and increased abundance of soil parasites and pathogens.

The global transport of nutrients has major consequences for adaptation and mitigation under climate change. For instance, both the loss of soil organic matter in feed-production regions and the shift to bacteria-based soil food webs in feed-consumption regions result in lowered soil stability (*14*). As a result, soils in both feed-production and feed-consumption locations will be more sensitive to climate change–induced extreme weather events. Restoring circularity by bringing back manure to feed-production areas is one option for addressing these imbalances, but that would increase global transport. Another possibility is to reduce the distance between feed production and consumption and close cycles within regions. Restoring circular food systems and therefore soil health will be a major challenge for enhancing sustainable food production and will require transdisciplinary research approaches that involve not only socioecology, economics, and other relevant disciplines but also a wide variety of stakeholders, practitioners, and policy-makers.

The decisions of farmers are influenced by numerous external factors, such as the price of land and the price of fertilizers, crop protection chemicals, and agricultural products on the world market. As a result, local soil conditions, including soil biodiversity and health, are influenced by national and global decisions, such as on trade and environmental protection. Soil health policy will also influence human decisions; for example, farmers may decide to move industrialized agricultural practices to less-industrialized countries. Thus, wherever enacted, soil health laws will need to account for these side effects to prevent soil health gain in one region from resulting in soil health loss elsewhere. Therefore, global networks of soil observatories—such as those initiated by SoilBON, LUCAS, and the Food and Agriculture Organization of the United Nations' International Network on Soil Biodiversity—need to be embedded into "living labs" that test experimental approaches under real-life conditions. In these living labs, researchers need to collaborate with all stakeholders to co-create knowledge that leads to sustainable development (*15*). At the core of such global approaches is the agreement between science and society on common definitions and gold standards of soil health.

PRESERVE WHAT IS HERE

It is essential that what is already here is protected and preserved. Draining carbon-rich soils, harvesting carbon from the remaining peatlands, or destroying pristine soil biodiversity to convert land for yield maximization of single crop species might generate short-term profit for some but are self-destructive for society as a whole in the long term. Moreover, climate change mitigation activities need to acknowledge the roles of soils in global biogeochemical cycles. Planting trees in natural grasslands and peatlands, for instance, can have undesirable consequences for biodiversity and the huge amounts of carbon in their soils. Real action is urgently needed to prevent vast amounts of fertile topsoil from being washed away in waterways and onto ocean floors. It takes only a few years to destroy what nature has built over centuries. Therefore, the protection of soils, soil biodiversity, and soil health should be high on the policy list of all nations and regions, because dead soil does not provide a sustainable business model anywhere. ■

REFERENCES AND NOTES

1. Food and Agriculture Organization of the United Nations (FAO), *et al*., "State of knowledge of soil biodiversity—Status, challenges and potentialities: Report 2020" (FAO, 2020); http://www.fao.org/documents/card/en/c/CB1928EN.
2. C. Folke *et al*., *Ambio* **50**, 834 (2021).
3. O. C. Ruppel, *Soil Secur* **6**, 100056 (2022).
4. European Academies Science Advisory Council, "Opportunities for soil sustainability in Europe" (EASAC Policy Report 36, German National Academy of Sciences Leopoldina, 2018), p. 48.
5. European Commission, "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. EU Soil Strategy for 2030: Reaping the benefits of healthy soils for people, food, nature and climate" (COM/2021/699 final European Commission, 2021); https://ec.europa.eu/environment/publications/eu-soil-strategy-2030_en.
6. J. Lehmann *et al*., *Nat. Rev. Earth Environ*. **1**, 544 (2020).
7. E. K. Bünemann *et al*., *Soil Biol. Biochem*. **120**, 105 (2018).
8. M. A. Tsiafouli *et al*., *Glob. Change Biol*. **21**, 973 (2015).
9. N.-P. Nguyen *et al*., *NPJ Biofilms Microbiomes* **2**, 16004 (2016).
10. C. A. Guerra *et al*., *Nat. Commun*. **11**, 3870 (2020).
11. R. D. Bardgett, W. H. van der Putten, *Nature* **515**, 505 (2014).
12. J. Fediajevaite *et al*., *Ecol. Evol*. **11**, 4803 (2021).
13. A. Orgiazzi *et al*., Global Soil Biodiversity Atlas (European Commission, Publications Office of the European Union, 2016); https://www.globalsoilbiodiversity.org/atlas-introduction.
14. F. T. de Vries *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14296 (2013).
15. C. A. Guerra *et al*., *Science* **371**, 239 (2021).

ACKNOWLEDGMENTS

All coauthors contributed equally to this work.

10.1126/science.abn7248

BOOKS *et al.*

Some actions, such as mitigating environmental pollution, will improve the lives of many animals.

PHOTO: ZANSKAR/ISTOCK PHOTO

ANIMAL STUDIES

The right to strive in a changing world

A philosopher prioritizes agency and Anthropocene concerns in a theory of animal justice

By **Dale Jamieson**

Since the publication of Peter Singer's *Animal Liberation* in 1975, attitudes toward nonhuman animals have changed drastically (*1*). Singer's book was followed by papers and monographs by other philosophers, and a new field of animal studies emerged in its wake. *Justice for Animals*, by the eminent philosopher and law professor Martha Nussbaum, is a welcome addition to this literature.

According to Nussbaum, all sentient animals have "imperfect rights"—rights not against any particular person but against "humanity imagined as capable of collective action." These rights, she argues, should be recognized as actionable legal rights.

The "Capabilities Approach," a theory of well-being that Nussbaum worked out in its original form with Nobel Prize winner Amartya Sen, provides the foundation for a substantive account of what it is for a creature to flourish and thus what opportunities rights should protect. The core idea is that an injustice occurs when an animal's "significant striving" is inhibited by "wrongful thwarting" either through negligent or deliberate action. Nussbaum argues for the superiority of her theory over those of others and goes on to discuss the harm of death, tragic conflicts, our responsibilities to animals in a variety of settings, the possibility of human friendship with animals, conflicts of duty, and the role of law in achieving justice for animals.

Nussbaum's writing is energetic and direct, full of stories and anecdotes. The book is pleasingly constructive rather than deadeningly polemical. Its major accomplishments, in my view, are the following: First, by showing how the Capabilities Approach can support a theory of justice for animals, it strengthens the argument that our systematic treatment of nonhuman animals is morally indefensible from almost any normative perspective.

The Capabilities Approach also puts respect for agency at the center of what we owe animals. Previous work has tended to neglect agency and focused instead on animal suffering (*2*). This is understandable given the sheer scale of suffering entailed by practices such as the mass production of meat. Yet the neglect of agency can obscure the full range of ways in which we encounter animals.

Justice for Animals: Our Collective Responsibility
Martha C. Nussbaum
Simon & Schuster, 2023. 400 pp.

Another strength of this book is Nussbaum's full-on embrace of the Anthropocene. She acknowledges that directly or indirectly, through action or omission, humans affect virtually every living thing on the planet, thus giving rise to broad and deep obligations. This is in welcome contrast to views that tend to cordon off "nature" and relieve humans of responsibility for what happens there. There are also valuable discussions of such topics as why death can be a harm and how extinction hurts individual animals, as well as interesting historical asides.

Nussbaum envisions a "multispecies world in which all have opportunities for flourishing." Although she acknowledges conflicts and challenges, there is an overall sense of unworldly optimism in this book. She understands how important human population control is to realizing her vision, for example, but seems to think that educating women is enough to put this concern to rest. She also discusses various problems that arise from keeping companion animals, but there is little mention of, for example, the toll that domestic cats take on songbirds. Nor does she really address the discomfort that many people will feel about establishing a government office of "domestic animal welfare" or allowing animals to sue their guardians.

Nussbaum tends to be too categorical in attributing commonsense psychological states to animals, and some of her most controversial claims (e.g., that fish can be painlessly killed) are not well anchored in the literature she cites. There is little acknowledgment of the constancy of change in nature and the disruption that it entails. Whether in the Anthropocene or some other epoch, natural selection will drive speciation and extinction, and nutrients in any form, whether bacterial or mammalian, are seldom left on the table. The multispecies world will never be a peaceable kingdom.

In a book of such sweep and ambition, marked by the clarity and boldness of its claims, there is bound to be a lot with which one can take issue. Indeed, it is a virtue of this book that the mind of the author is so clearly on display. Even if in the end they are not fully convinced by Nussbaum's arguments, a confirmed speciesist or a Benthamite utilitarian can read this book with pleasure and profit. ■

The reviewer is at the Center for Environmental and Animal Protection, New York University, New York, NY 10003, USA. Email: dale.jamieson@nyu.edu

REFERENCES AND NOTES

1. P. Singer, *Animal Liberation* (HarperCollins, 1975).
2. D. Jamieson, *Harv. Rev. Philos.* **25**, 111 (2018).

10.1126/science.adf6029

SCIENCE AND SOCIETY

Materialism meets transcendence

A new documentary series follows a physicist's pursuit of the profound

By **Robert S. Krauss**

Some years ago on a summer night, Massachusetts Institute of Technology physicist, philosopher, and novelist Alan Lightman lay down in a boat off a remote coast of Maine, stared up at the stars, and experienced a transcendental feeling of deep connection to nature and the cosmos. What had happened in that moment? he later wondered. This question sent Lightman on a personal journey in which he attempted to reconcile his physicist's materialism with his profound experience in that boat. Lightman now shares his journey in *Searching: Our Quest for Meaning in the Age of Science*.

Alan Lightman's efforts to answer life's biggest questions reveal the joy of scientific searching.

In part 1, "The Stars & The Osprey," Lightman undergoes functional magnetic resonance imaging and interviews neuroscientist Robert Desimone about how much neuroimaging can tell us about Lightman's transcendental experience. He ultimately finds this approach unsatisfying and introduces viewers to the debate between mechanists, who believe that the laws of physics, chemistry, and biology are sufficient to explain life, and vitalists, who believe that living creatures are imbued with an additional spiritual quality not explainable by science.

Here, he interviews biologist and Nobel laureate Jack Szostak and the Dalai Lama as proponents of these two camps, respectively. Although the Dalai Lama expresses enthusiasm for scientific investigation, most scientists will likely resonate with Szostak's declaration that "It's not *just* atoms and molecules, it's the organization…it's no less wonderful or beautiful because we understand that there is a natural origin for [life]."

Part 2, "The Big & The Small," begins with the familiar "powers of 10" view of the Universe, moving from the subatomic to the galactic. Lightman then queries what such explorations have to do with consciousness, conversing with BINA48, an extraordinary humanoid robot programmed by artificial intelligence with >100 hours of a real woman's memories. Here, he speculates that BINAs of the future may achieve consciousness. He then probes this issue over Zoom with the Dalai Lama, whom viewers observe watching a movie of BINA48 conversing with the woman from whom BINA48 was programmed; this multilayered interaction is simultaneously disconcerting, comical, and wondrous. After additional interviews with a bioethicist, a rabbi, and others, Lightman ultimately concludes that we may be just atoms and molecules, but, in the words of Emily Dickinson, "The brain is wider than the sky."

In part 3, "Homo Techno," Lightman contemplates our inner world of perception, consciousness, and self-awareness. He meets Erik Sorto, who lost all movement from the neck down after being shot in the back. With electrodes implanted into his posterior parietal cortex and 2 years of training, Sorto learned to control the movement of a robotic arm by thought. This is an extraordinary achievement, and Lightman posits that it is an example of the beginning of our transition from *Homo sapiens* into *Homo techno*, part-human, part-machine entities that reflect the modification of human evolution by technological means. From an actual evolutionary biology standpoint, this is nonsensical, and it is unclear that Lightman even means to propose such an idea, but this section's loose language will rankle some viewers.

At another point in the series' final episode, Lightman finds himself dizzy from talk of neurons and galaxies and takes refuge in closely examining a single square inch of earth. His biophilia is obviously meaningful to him, and it would have been stimulating had the episode included interviews with an evolutionary biologist or naturalist, who might have helped to better articulate this facet of the human experience. The series closes, appropriately, with philosophical musings about the need for each of us to find meaning for ourselves.

A small problem at the outset is the inherent impossibility of conveying transcendence through description—the degree to which viewers relate to Lightman's moment of enlightenment will depend on their own experiences and inclinations. Additionally, Lightman's screen persona leaves something to be desired. There are, however, few people better qualified to explore these issues, and as the series progresses, his humanity shines through, bringing a welcome lightness to some potentially ponderous material.

Despite its focus on phenomena currently unexplainable by science, *Searching* is full of the joy and passion that can be found in the doing of science and succeeds in conveying how deeply meaningful science is to its practitioners. It is well worth your time and is especially recommended to families with kids curious about life and our world. ■

Searching: Our Quest for Meaning in the Age of Science
Geoffrey Haines-Stiles, director
Premieres 7 January 2023 on PBS.org and on public television stations; check local listings.

The reviewer is at the Department of Cell, Developmental, and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. Email: robert.krauss@mssm.edu

10.1126/science.adf7352

PHOTO: *SEARCHING: OUR QUEST FOR MEANING IN THE AGE OF SCIENCE*

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PLANT SCIENCE

Diversity reveals infection resistance

Parasitic witchweed (*Striga*) reduces the yield of maize grown in infected fields. Strigolactones from maize roots encourage *Striga* germination. Li *et al.* analyzed the natural variation in types of strigolactones exuded from maize roots. Maize genotypes that produced mainly zealactol suffered less *Striga* infection than those that produced mainly zealactone. A single cytochrome P450 catalyzes several of the oxidative steps in strigolactone biosynthesis, including conversion of precursors to either zealactol or zealactone. —PJH *Science*, abq4775, this issue p. 94

A flowering *Striga* plant overwhelming a susceptible maize host

PHOTO: ETHAN BASS, CREATIVE COMMONS SHARE ALIKE 4.0

BRAIN ANATOMY

An extra layer lines the brain

The traditional view is that the brain is surrounded by three layers, the dura, arachnoid, and pia mater. Møllgård *et al*. found a fourth meningeal layer called the subarachnoid lymphatic-like membrane (SLYM). SLYM is immunophenotypically distinct from the other meningeal layers in the human and mouse brain and represents a tight barrier for solutes of more than 3 kilodaltons, effectively subdividing the subarachnoid space into two different compartments. SLYM is the host for a large population of myeloid cells, the number of which increases in response to inflammation and aging, so this layer represents an innate immune niche ideally positioned to surveil the cerebrospinal fluid. —SMH

Science, adc8810, this issue p. 84

MPOX

The mpox virus DNA-synthesizing machine

Mpox (formerly called monkeypox) is an active public health emergency that requires more effective means of treatment and prevention. Peng *et al*. focused on a complex that plays a key role in the genome replication process of the mpox virus. They used cryo–electron microscopy to determine a structure of the DNA polymerase F8, which catalyzes the synthesis of viral DNA, in complex with the processivity factor consisting of A22 and E4 and the DNA substrate. The mode of operation of the processivity factor shown by the structure differs from existing ideas and may provide a basis for antiviral drug design. —DJ

Science, ade6360, this issue p. 100

EXPATRIATE SCHOLARS

Measuring returning scientists' success

China is a top sender of students overseas, and the Chinese government launched the Young Thousand Talents program to recruit and nurture high-caliber, early-career expatriate scientists who return to China after they receive doctorates abroad. Shi *et al*. examined how effective the program has been in supporting the young scholars' productivity when they return to China compared with their peers that remained overseas. They found that the scholars were high (but not top) caliber and outperformed overseas peers in last-authored publications because of greater access to larger research teams and better research funding in China. —EEU

Science, abq1218, this issue p. 62

INNATE IMMUNITY

Keeping flagellin silent

Flagellin is a protein in the bacterial flagellum that can stimulate Toll-like receptor 5 (TLR5)–mediated immune responses; however, some commensal bacteria manage to avoid triggering immune activation. By screening publicly available gut metagenomes, Clasen *et al*. identified a class of "silent" flagellins that bind to TLR5 but only induce weak activation. The activity of TLR5 is controlled by a domain that is absent in these silent flagellins. Such proteins were found in human organoids, mice, and human stool and were produced by *Lachnospiraceae* gut bacteria. —DAE

Sci. Immunol. **8**, eabq7001 (2023).

METALLURGY

Tracking down the pores

Laser fusion techniques build metal parts through a high-energy melting process that too often creates structural defects in the form of pores. Ren *et al*. used x-rays to track the formation of these pores while also making observations with a thermal imaging system. This setup allowed the authors to develop a high-accuracy method for detecting pore formation from that thermal signature with the help of a machine learning method. Implementing this sort of tracking of pore formation would help avoid building parts with high porosity that are more likely to fail. —BG

Science, add4667, this issue p. 89

DEVELOPMENT

Going with the flow

In most vertebrates, left-right differences are specified during early embryogenesis by a small cluster of cells called the left-right organizer. Within this organizer, motile cilia move rapidly to create a leftward directional flow of extracellular fluid that is the first sign of a left-right difference, but how this flow is sensed and transduced into later molecular and anatomical left-right asymmetry has been unclear. Working with mouse embryos, Katoh *et al*. found that immotile cilia sense the mechanical force generated by the flow and suggest a biophysical mechanism by which the direction of the flow is sensed. Independently, working in zebrafish, Djenoune *et al*. used optical tweezers and live imaging to show that immotile cilia in the organizer function as mechanosensors that translate extracellular fluid flow into calcium signals. When motile cilia were paralyzed and normal flow stopped, mechanical manipulation of the cilia could rescue, or even reverse, left-right patterning. Thus, ciliary force sensing is necessary, sufficient, and instructive for embryonic laterality. —SMH

Science, abq8148, abq7317, this issue p. 66, p. 71

CORONAVIRUS

Protection spanning viral variants

Despite the success of the early generations of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccines, viral evolution and immune evasion have made the development of next-generation broadly protective vaccines essential. Zhao *et al*. used the evolutionary history of the SARS-CoV-2 S protein to develop S_{pan}, a vaccine antigen that carries amino acid residues that are consistent across SARS-CoV-2 strains. When administered to mice, S_{pan} elicited a more broadly neutralizing antibody response than a wild-type S protein vaccine. S_{pan} also conferred substantial protection against challenge with the Beta, Delta, and Omicron variants despite Delta and Omicron arising after S_{pan} was designed. —CSM

Sci. Transl. Med. **15**, eabo3332 (2023).

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

VACCINES

Mapping vaccine-induced immune signatures

Most studies examining vaccine-induced immune responses have largely focused on one vaccine at a time, which provides a limited understanding of the advantages and disadvantages of various formulations. Looking for a more systemic overview, Hagan *et al.* analyzed the human transcriptional responses to 13 vaccines against various pathogens from previously published datasets. Many of these vaccines triggered expression from some common gene signatures linked to the induction of innate immunity and plasmablasts, although the effect of each vaccine was different. Once the authors controlled for the magnitude and kinetics of each vaccine's distinct immune response, they were able to use machine learning to develop a common time-adjusted signature that could predict antibody responses to the vaccines. —STS

Nat. Immunol. **23**, 1788 (2022).

CANCER

Make your own neoantigens

Immunotherapies activate the body's own immune system to fight cancer, but their effectiveness is limited by their ability to distinguish cancer cells from healthy cells. Chemical inhibitors are a mainstay of cancer therapy, but they too have limitations because cancer cells frequently become resistant to drug treatment. Putting the two together, Hattori *et al*. developed an approach that draws on the strengths of both drug therapy and immunotherapy. The authors first treated lung cancer with the drug sotorasib, which targets *KRAS*, a mutant gene involved in cell signaling pathways that control cell growth, cell maturation, and cell death. Engineered antibodies were then administered that specifically recognize neoantigens created by mutant KRAS protein bound to sotorasib. The antibodies killed even

PHOTO: PATRICK AVENTURIER/GETTY IMAGES

VITICULTURE

Climate change shifts wine country

Among other threats to biodiversity and livelihoods, climate change is predicted to have major effects on European wine production. Sgubin *et al.* modeled changes to phenological phases, including flowering and fruit maturity, for early- to late-ripening grape varieties under two global warming scenarios. With increasing temperatures, potential wine-producing areas are expected to expand north across Europe. Swaths of traditional wine-producing areas will likely become unsuitable for grapevines by 2100. Predicted losses will occur at much higher rates if warming exceeds 2°C, with limited ability to compensate by switching wine varietals. Extreme heating or cooling events may alter predictions, but maintaining future warming below 2°C could help to ensure the future production of high-quality wine. —BEL

Glob. Chang. Biol. 10.1111/gcb.16493 (2022).

Climate change is threatening traditional wine production areas such as this drought-stricken site in southern France.

later in life among the children before they gave birth to the grandchildren. —BW

Am. Econ. Rev. 10.1257/aer.20210937 (2022).

THERMOELECTICS

Tuning the transport

Thermoelectric materials can convert heat into electricity, and many of these materials are becoming increasingly competitive with bismuth telluride for commercial applications. Improving thermoelectric efficiency usually requires decoupling the electrical and thermal transport properties. Yu *et al.* discovered that planar defects called quantum gaps manage this decoupling in germanium telluride. The presence of these defects results in a material with improved thermoelectric properties. Engineering quantum gaps into other thermoelectric materials should help to improve their efficiencies. —BG

Nat. Commun. **13**, 5612 (2022).

sotorasib-resistant cancer cells but spared normal cells that do not have mutant KRAS. —YN

Cancer Discov. 10.1158/2159-8290.CD-22-1074 (2022).

ENERGY STORAGE

Thermal energy storage in salt hydrates

Most of the energy produced in the United States ends up wasted as heat, one of the major challenges for our sustainable future. Thermal energy storage (TES) technologies aim to minimize these losses. Among various TES systems, those based on thermochemical materials (TMs) demonstrate the highest energy densities, but they have not been actively studied so far. Using density functional theory and machine learning models, Kiyabu *et al.* developed an efficient approach to search for new TMs among salt hydrates, a promising class of TMs, and to establish design guidelines for maximizing their energy density. The authors identified a series of hydrates that demonstrate operating temperatures suitable for practical applications, with energy densities surpassing the US Department of Energy's technical target for TES. —YS

J. Am. Chem. Soc. **144**, 21617 (2022).

HEALTH ECONOMICS

Health care benefits span generations

Expanded access to health care for low-income pregnant women not only improved the health of their newborn children, but also contributed, years later, to improved health of their newborn grandchildren. Analyzing records on all US births from 1975 to 2017 to study expansion of the Medicaid program in the 1980s, East *et al.* suggest that impacts on the grandchildren of improved in utero care for the children were not driven by changes in fertility behavior or selection into pregnancy as the children grew up. Instead, expanded prenatal care likely contributed to better health and economic conditions

Expanded health care for low-income pregnant women resulted years later in the improved health of their newborn grandchildren.

CELL BIOLOGY

Fixing leaky lysosomes

Lysosomes are a key degradative compartment within cells, so their integrity is essential. A membrane repair process involving the so-called ESCRT machinery is known to help maintain lysosomal integrity by pinching out damaged regions of membrane. Radulovic *et al.* describe an additional repair mechanism that involves recruitment of the endoplasmic reticulum (ER) and cholesterol transfer. Lipidomic analyses of isolated lysosomes after membrane damage showed increased lysosomal phosphatidylserine, cholesterol, and phosphorylated phosphatidylinositol species. The ER proteins VAPA and VAPB and the cholesterol-binding protein ORP1L were recruited to damaged lysosomes and thereby provided a source of membrane for successful repair. Interference with this process resulted in cell death after lysosomal membrane damage. —SMH

EMBO J. **41**, e112677 (2022).

PHOTO: DSZC/GETTY IMAGES

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

EPIDEMIOLOGY

Enabling wastewater surveillance

The COVID-19 pandemic has highlighted the importance and effectiveness of analyzing wastewater to understand disease spread, prevalence, and evolution. In a Perspective, Levy *et al.* discuss different approaches to wastewater surveillance that can enable pathogen detection at the community level, as well as at the level of individual buildings. Additionally, wastewater analyses can elucidate the emergence and prevalence of new variants, which can inform public health services. There are also untargeted approaches to monitor known pathogens and inform outbreak risk. Although further technical advances are needed to improve the deployment and uptake of wastewater surveillance across the globe, wastewater is an important resource that has much to offer public health monitoring. —GKA

Science, ade2503, this issue p. 26

OPTICS

Lithium niobate photonics

The optoelectronic and nonlinear optical properties of lithium niobate make it a workhorse material for applications in optics and communication technology. Boes *et al*. reviewed the science and technology of lithium niobate and its role in various aspects of photonic technology. They surveyed the evolution from bulk lithium niobate through weakly confining waveguides to the recent developments with thin-film lithium niobate. The ability to span the entire spectral range from radio to optical wavelengths illustrates the versatility of lithium niobate as a platform material in integrated photonics. —ISO

Science, abj4396, this issue p. 40

BIOCHEMISTRY

Mutant misbehavior at the membrane

Dynamic remodeling of the neuronal plasma membrane is critical for neuronal development. Beeman *et al*. investigated why mutations in the kinase TAOK1 are associated with neurodevelopmental disorders. TAOK1 bound to plasma membrane phospholipids through a helical region, and binding was inhibited by phosphorylation of that region by the kinase domain. Kinase domain mutations associated with neurodevelopmental disorders abrogated catalytic activity, locking TAOK1 into a membrane-bound state that caused abnormal protrusions and growth and maturation defects in dendrites of mouse hippocampal neurons. —LKF

Sci. Signal. **16**, eadd3269 (2023).

GLACIERS

Melting away

Mountain glaciers, perennial ice masses excluding the Greenland and Antarctic ice sheets, are a critical water resource for nearly two billion people and are threatened by global warming. Rounce *et al*. projected how those glaciers will be affected under global temperature increases of 1.5° to 4°C, finding losses of one quarter to nearly one half of their mass by 2100 (see the Perspective by Aðalgeirsdóttir and James). Their calculations suggest that glaciers will lose substantially more mass and contribute more to sea level rise than current estimates indicate. —HJS

Science, abo1324, this issue p. 78; see also ade2355, p. 29

OPTICS

Determining limits on optical thickness

The functionality and performance of optical components has largely been limited by the optical materials available. Recent developments in fabrication and optical component design now widen what can be achieved. In particular, flat optics or metasurfaces can be designed with the functionality of bulk optical components but with their thickness shrunk to just several hundred nanometers. Miller now presents a theoretical study to determine what the minimum thickness can be for a specific optical function (see the Perspective by Monticone). The approach is general and should provide bounds on the minimum size of other wave systems, including radio and acoustic systems. —ISO

Science, ade3395, this issue p. 41; see also adf2197, p. 30

IMMUNOLOGY

Lingering immune changes after obesity

A past period of obesity caused by a high-fat diet in mice produces persistent changes in innate immunity even after weight loss and normalization of metabolism. Hata *et al*. found that such diet-induced obesity in mice, even after it was resolved, led to persistent epigenetic changes in chromatin in macrophages associated with increased expression of genes that function in inflammatory responses (see the Perspective by Mangum and Gallagher). Experiments with transplants of adipose tissue or bone marrow implicated alterations of myeloid cells in exacerbating inflammatory responses to experimentally induced injury in the eye. If similar processes occur in humans, then the authors propose that such changes could contribute to predisposition to age-related macular degeneration associated with obesity. —LBR

Science, abj8894, this issue p. 45; see also adf6582, p. 28

Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **ScienceAdvances.org**

Science Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS

REVIEW SUMMARY

OPTICS

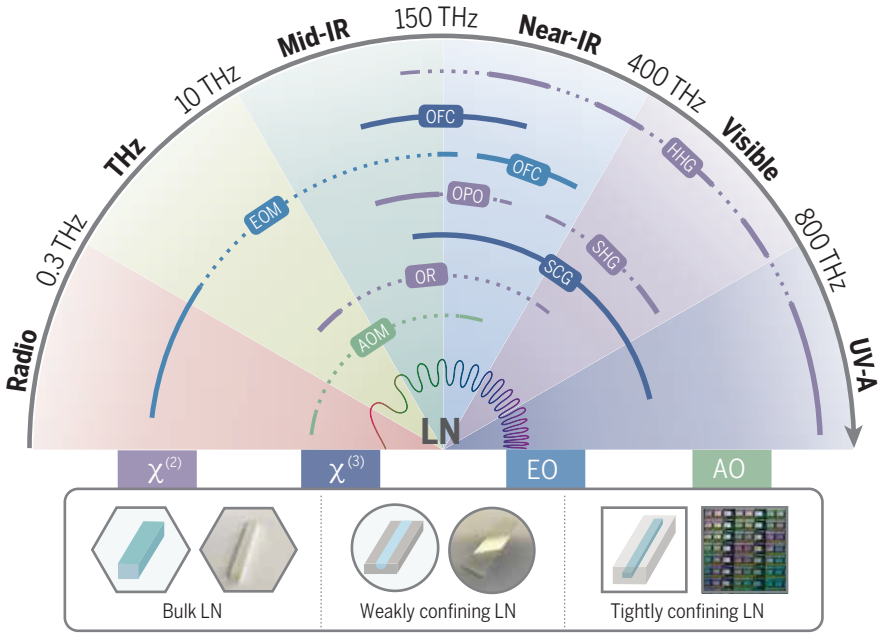
Lithium niobate photonics: Unlocking the electromagnetic spectrum

Andreas Boes*†, Lin Chang*†, Carsten Langrock, Mengjie Yu, Mian Zhang, Qiang Lin, Marko Lončar, Martin Fejer, John Bowers, Arnan Mitchell

BACKGROUND: Electromagnetic (EM) waves underpin modern society in profound ways. They are used to carry information, enabling broadcast radio and television, mobile telecommunications, and ubiquitous access to data networks through Wi-Fi and form the backbone of our modern broadband internet through optical fibers. In fundamental physics, EM waves serve as an invaluable tool to probe objects from cosmic to atomic scales. For example, the Laser Interferometer Gravitational-Wave Observatory and atomic clocks, which are some of the most precise human-made instruments in the world, rely on EM waves to reach unprecedented accuracies.

This has motivated decades of research to develop coherent EM sources over broad spectral ranges with impressive results: Frequencies in the range of tens of gigahertz (radio and microwave regimes) can readily be generated by electronic oscillators. Resonant tunneling diodes enable the generation of millimeter (mm) and terahertz (THz) waves, which span from tens of gigahertz to a few terahertz. At even higher frequencies, up to the petahertz level, which are usually defined as optical frequencies, coherent waves can be generated by solid-state and gas lasers. However, these approaches often suffer from narrow spectral bandwidths, because they usually rely on well-defined energy states of specific materials, which results in a rather limited spectral coverage.

To overcome this limitation, nonlinear frequency-mixing strategies have been developed. These approaches shift the complexity from the EM source to nonresonant-based material effects. Particularly in the optical regime, a wealth of materials exist that support effects that are suitable for frequency mixing. Over the past two decades, the idea of manipulating these materials to form guiding structures (waveguides) has provided improvements in efficiency, miniaturization, and production scale and cost and has been widely implemented for diverse applications.



Lithium niobate spectral coverage. The EM spectral range and processes for generating EM frequencies when using lithium niobate (LN) for frequency mixing. AO, acousto-optic; AOM, acousto-optic modulation; $\chi^{(2)}$, second-order nonlinearity; $\chi^{(3)}$, third-order nonlinearity; EO, electro-optic; EOM, electro-optic modulation; HHG, high-harmonic generation; IR, infrared; OFC, optical frequency comb; OPO, optical paramedic oscillator; OR, optical rectification; SCG, supercontinuum generation; SHG, second-harmonic generation; UV, ultraviolet.

ADVANCES: Lithium niobate, a crystal that was first grown in 1949, is a particularly attractive photonic material for frequency mixing because of its favorable material properties. Bulk lithium niobate crystals and weakly confining waveguides have been used for decades for accessing different parts of the EM spectrum, from gigahertz to petahertz frequencies. Now, this material is experiencing renewed interest owing to the commercial availability of thin-film lithium niobate (TFLN). This integrated photonic material platform enables tight mode confinement, which results in frequency-mixing efficiency improvements by orders of magnitude while at the same time offering additional degrees of freedom for engineering the optical properties by using approaches such as dispersion engineering. Importantly, the large refractive index contrast of TFLN enables, for the first time, the realization of lithium niobate–based photonic integrated circuits on a wafer scale.

OUTLOOK: The broad spectral coverage, ultralow power requirements, and flexibilities of lithium niobate photonics in EM wave generation provides a large toolset to explore new device functionalities. Furthermore, the adoption of lithium niobate–integrated photonics in foundries is a promising approach to miniaturize essential bench-top optical systems using wafer scale production. Heterogeneous integration of active materials with lithium niobate has the potential to create integrated photonic circuits with rich functionalities. Applications such as high-speed communications, scalable quantum computing, artificial intelligence and neuromorphic computing, and compact optical clocks for satellites and precision sensing are expected to particularly benefit from these advances and provide a wealth of opportunities for commercial exploration. Also, bulk crystals and weakly confining waveguides in lithium niobate are expected to keep playing a crucial role in the near future because of their advantages in high-power and loss-sensitive quantum optics applications. As such, lithium niobate photonics holds great promise for unlocking the EM spectrum and reshaping information technologies for our society in the future. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: andy.boes@adelaide.edu.au (A.B.); linchang@pku.edu.cn (L.C.)
†These authors contributed equally to this work.
Cite this article as A. Boes *et al.*, *Science* **379**, eabj4396 (2023). DOI: 10.1126/science.abj4396

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abj4396

REVIEW

OPTICS

Lithium niobate photonics: Unlocking the electromagnetic spectrum

Andreas Boes[1,2,3]*†, Lin Chang[4,5]*†, Carsten Langrock[6]‡, Mengjie Yu[7,8], Mian Zhang[9], Qiang Lin[10], Marko Lončar[7], Martin Fejer[6], John Bowers[11], Arnan Mitchell[1]

Lithium niobate (LN), first synthesized 70 years ago, has been widely used in diverse applications ranging from communications to quantum optics. These high-volume commercial applications have provided the economic means to establish a mature manufacturing and processing industry for high-quality LN crystals and wafers. Breakthrough science demonstrations to commercial products have been achieved owing to the ability of LN to generate and manipulate electromagnetic waves across a broad spectrum, from microwave to ultraviolet frequencies. Here, we provide a high-level Review of the history of LN as an optical material, its different photonic platforms, engineering concepts, spectral coverage, and essential applications before providing an outlook for the future of LN.

The technological development of our modern society is closely linked to our ability to make use of electromagnetic (EM) waves. The wide EM spectrum, spanning from radiowaves and microwaves through infrared radiation, visible light, and ultraviolet (UV) radiation up to high-energy x- and γ-rays, has transformed the way we record images, carry information, and transmit energy. Driven by the sophisticated control of EM waves, the past few decades have witnessed notable breakthroughs in a wide range of areas such as high-speed communication (*1*, *2*), ultraprecision time-frequency metrology (*3*–*5*), bioimaging (*6*–*8*) and quantum-information science (*9*, *10*).

The generation and manipulation of EM waves lies at the heart of all scientific and technological explorations. Depending on the frequencies, there are several main strategies for generation and processing: Radio frequency (RF) (<~100 GHz) signals can readily be produced by microwave oscillators and then manipulated through conventional complementary metal-oxide-semiconductor (CMOS) electronics. Signals with frequency greater than 100 GHz (typically referred to as millimeter waves) up to a few THz can be generated, for example, by resonant tunneling diodes and processed by high-speed electronics that use silicon or III-V semiconductors. For higher frequencies (10 THz to 1 PHz) in the optical regime, the most common methods for EM wave generation use solid-state, fiber, gas, and semiconductor lasers. Although these strategies have successfully enabled many applications, each platform individually can only offer a limited spectral coverage because of the specific platform's dependence dependence on well-defined energy bands or levels of solid-state materials, atoms, and molecules.

To overcome this limitation, another important strategy, parametric nonlinear frequency mixing, was introduced. Starting from optical frequencies (hundreds of THz), this process leverages broadband parametric nonlinear effects for multiwave mixing to unlock previously inaccessible EM frequencies on demand. This approach shifts the complexity from a custom EM wave source to nonresonant material effects, which can be engineered with additional degrees of freedom, allowing access to a much broader part of the EM spectrum (THz to PHz) with unprecedented control and performance. In addition to nonlinear frequency mixing, the GHz to THz part of the EM spectrum can be bridged through microwave-to-optical conversion, enabling the efficient processing of EM waves through well-established CMOS electronics.

These prospects spurred the development of a wide range of nonlinear and electro-optic material platforms over the past few decades. Among these platforms, lithium niobate ($LiNbO_3$, or LN), which has been described as the "the silicon of photonics" (*11*), turns out to be particularly suitable for the generation and manipulation of EM frequencies because it offers a rare combination of advantageous properties: (i) large electro-optic, piezoelectric, and nonlinear-optic material coefficients; (ii) engineerability of velocity matching through quasi–phase matching (QPM) and waveguide dispersion; (iii) broad transparency (400 nm to 5 μm); (iv) long-term stability; and (v) widespread commercial availability of large, low-cost, optical-quality wafers. This makes LN one of the key photonic materials that has the potential to expand access to an ultrawide part of the EM spectrum and support the next generation of scientific breakthroughs and commercial products.

LN platforms

LN is a ferroelectric crystal and was first synthesized in 1949 (*12*) in its polycrystalline form. From this discovery, it took 15 years until further studies identified the material's characteristic electro-optical (*13*) and second-order nonlinear-optical (*14*) properties. The growth of single-crystalline LN using the Czochralski technique (*15*, *16*) represented a breakthrough; this technique is still in use today and is able to produce optical-quality wafers up to a diameter of 150 mm (6 inches), and several crystal compositions—such as congruent, near-stoichiometric, or doped with alkaline or transition metals—are available commercially. Over the decades, three main LN photonic platforms have emerged, namely bulk crystals, weakly confining waveguides, and tightly confining waveguides, whose evolution can be found in Fig. 1.

Bulk LN crystals

Bulk LN crystals have found wide adoption for generation and manipulation of EM waves owing to their compatibility with free-space optical setups, ability to handle high optical power, ease of fabrication, and low cost. Such crystals are typically millimeter- to centimeter-scale blocks of LN with optical-grade polished facets (see right side of Fig. 1).

Early demonstrations in bulk LN crystals include electro-optic modulation (*13*) and second-harmonic generation (SHG) (*14*). A breakthrough discovery occurred when it was observed that the spontaneous polarization of LN crystals could be inverted locally by applying a high electric field at room temperature (*17*). This process, referred to as "electric-field poling," opened a reliable path for engineering the phase-velocity matching (i.e., momentum conservation) between different waves and made previously explored domain-inversion methods that relied on high-temperature ionic diffusion processes (*18*) obsolete. Photorefraction was first discovered when investigating LN for nonlinear devices (*19*), which later provided the means for high-density data storage in LN (*20*).

[1]Integrated Photonics and Applications Centre (InPAC), School of Engineering, RMIT University, Melbourne, VIC 3000, Australia. [2]Institute for Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide, SA 5005, Australia. [3]School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia. [4]State Key Laboratory of Advanced Optical Communications System and Networks, School of Electronics, Peking University, Beijing 100871, China. [5]Frontiers Science Center for Nano-optoelectronics, Peking University, Beijing 100871, China. [6]Edward L. Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA. [7]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [8]Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, USA. [9]HyperLight, Cambridge, MA 02139, USA. [10]Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627, USA. [11]Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA.
***Corresponding author. Email: andy.boes@adelaide.edu.au (A.B.); linchang@pku.edu.cn (L.C.)** †These authors contributed equally to this work. ‡Present address: Keysight Laboratories, Santa Clara, CA 95952, USA.

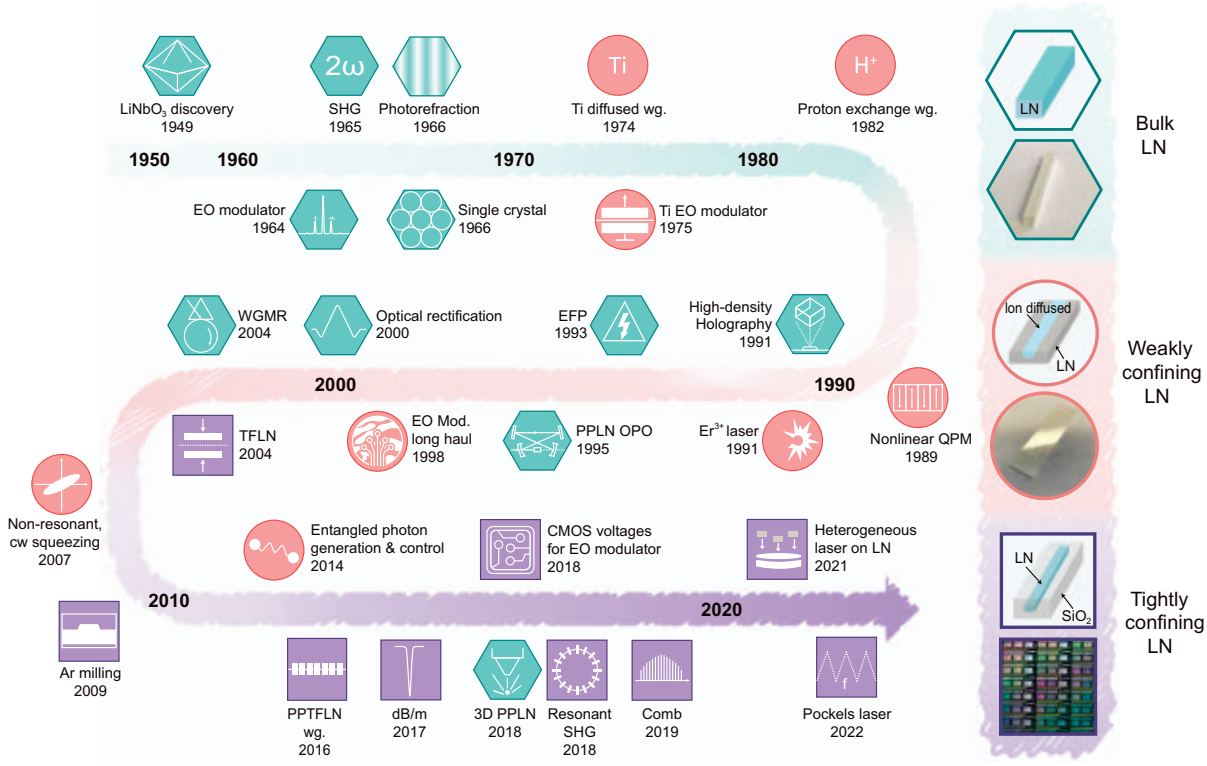


Fig. 1. Timeline of LN as a photonic material. LN has been developed into three major platforms: bulk crystals, weakly confining waveguides, and tightly confining waveguides (indicated by blue, pink, and purple, respectively). Milestone demonstrations are highlighted. cw, continuous-wave; EO, electro-optic; wg., waveguide; EFP, electric-field poling; WGMR, whispering gallery resonator; PPTFLN, periodically poled thin film LN.

Bulk crystals have also been formed in the shape of discs or ring cavities, by careful polishing of their facets (*21*–*23*). Such discs can form so-called whispering gallery resonators (*24*), with quality factors reaching hundreds of millions (*25*), making them attractive for highly coherent optical wave or microwave sources and nonlinear-optical applications. Recently, it was shown that ferroelectric domain engineering can also be achieved in three dimensions by using femtosecond laser pulses that are focused into the crystal (*26*). This demonstration opens opportunities for a new class of wave-mixing devices that were not previously feasible, such as three-dimensional nonlinear photonic crystals.

Bulk crystals are particularly attractive in optical cavity configurations such as parametric oscillators (*27*, *28*) (see Fig. 2) to enhance the nonlinear interaction. They are also attractive for high-power applications, for use inside laser cavities (Q-switch, intracavity SHG), or when using ultrashort, high–peak power laser pulses.

Weakly confining LN waveguides

Weakly confining LN waveguides maintain interacting fields over centimeter-length distances in small-mode volumes at high intensities (Fig. 1), thereby increasing nonlinear mixing efficiencies by two to three orders of magnitude when compared with bulk crystals. This relaxes the optical power requirements and enables efficient EM wave generation at moderate optical powers in the range of milliwatts (continuous-wave) or few nanojoules (pulsed).

Weakly confining LN waveguides can be formed by slightly altering the material composition or structure to locally increase the refractive index to form the guiding core. The first weakly confining waveguides were demonstrated by lithium out-diffusion (*29*), which was followed shortly thereafter by titanium in-diffusion (*30*) and later proton exchange (*31*) and femtosecond laser writing (*32*). Titanium in-diffusion and proton exchange remain common fabrication methods and require increased temperatures to drive the diffusion processes (~1100°C for titanium in-diffusion and ~200°C for proton exchange), which increase the extraordinary refractive index of LN by a few times 10^{-3} (*33*) and ~0.1 (*34*), respectively. To form low-loss, nonlinear-active waveguides through proton exchange, an annealing step at ~300°C is used, which reduces the index contrast to ~0.02. Such waveguides have been used for a number of frequency-mixing demonstrations such as the first implementation of QPM (see Fig. 3) for SHG in LN (*35*, *36*) and integrated erbium lasers (*37*). Importantly, when used as low-loss, high-speed electro-optic modulators, such waveguides were a key component for long-haul communication systems (*38*). This platform has been explored for low-loss quantum-optical applications, for example, up-conversion to near-visible wavelengths for single-photon detection (*39*) and on-chip entangled photon-pair generation by spontaneous parametric down-conversion and control of the generated photons (*40*). A large body of work in this field is reviewed in Gil-Lopez *et al.* (*41*).

Weakly confining waveguides are attractive because their mode volume is naturally close to that of standard optical fibers, which enables low interface losses (<0.5 dB at near-infrared wavelengths) between waveguide and fiber. Thus, a wide range of optical equipment that has been developed for high-speed optical communication can readily be connected to these waveguides, providing a large range of linear and nonlinear optical signal processing functionalities (*42*). These waveguides can be inexpensively produced on a wafer scale and only require standard lithographic tools with readily available micrometer resolution.

Tightly confining LN waveguides

Tightly confining LN waveguides are a relatively new class of LN structures with even smaller mode volumes that reach subwavelength mode

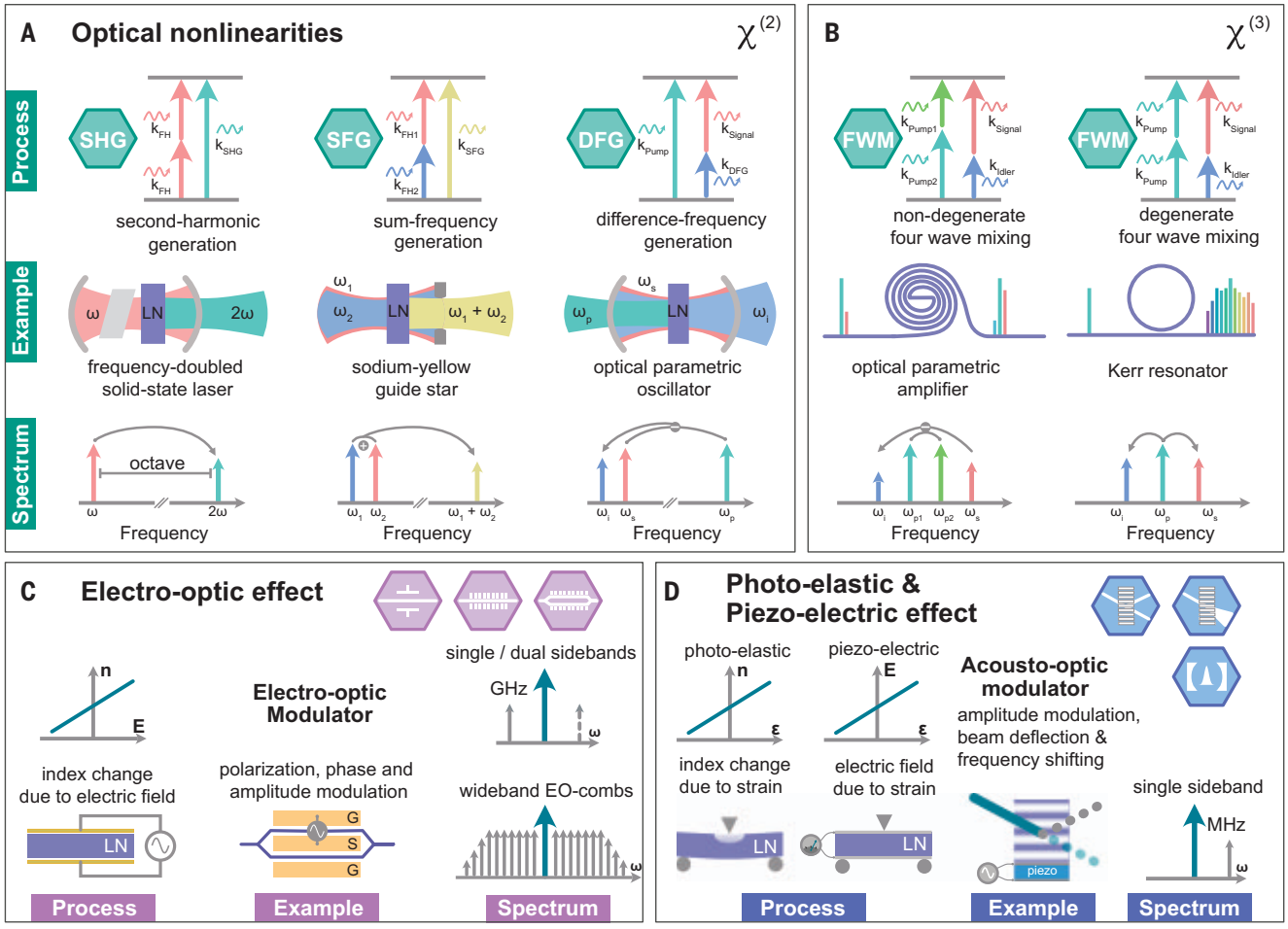


Fig. 2. LN material properties used for generating and manipulating EM waves. (**A** to **D**) Illustration of the (A) second-order and (B) third-order nonlinear-optic, (C) electro-optic, and (D) photo-elastic and piezo-electric material properties of LN that are used for efficient generation and manipulation of EM frequencies on demand.

diameters. This results in frequency mixing efficiencies that are nearly two orders of magnitude higher than those of weakly confining waveguides, in addition to offering broad integration and dispersion-engineering opportunities.

Such strongly confining waveguides (*43*) use a thin film of LN on a lower-index cladding layer, akin to the silicon-on-insulator platform, which can be manufactured at scale with good film uniformity and quality. Optical waveguides are typically made by dry etching (*44*) the LN film to form ridge waveguides or by strip-loading with a material that has a refractive index that is higher than that of the top cladding (*45*); however, methods such as laser ablation and diamond-blade scoring have also been explored. The optical modes in such waveguides are tightly confined because of the subwavelength waveguide dimensions and the high refractive index contrast between the guiding core and cladding, which enables dense integration using low-loss small-radius bends (*46*). This high index contrast has led to highly efficient and compact frequency-mixing and frequency-generating components such as frequency doublers (*45*, *47*, *48*), electro-optic modulators (*49*, *50*), optical frequency comb (OFC) generators (*51*–*53*), and lasers and amplifiers based on Er doping (*54*, *55*).

The thin-film LN (TFLN) platform enables photonic integration at a scale and density approaching that available on semiconductor platforms (*56*). This is a highly attractive proposition because the material properties of LN allow for the homogeneous integration of important active and passive photonic functions, often eliminating the need for additional materials. To incorporate lasers and detectors on TFLN, building blocks based on III-V semiconductor materials still require heterogeneous integration (*57*), which may be accomplished using techniques developed for other popular integrated-photonic platforms such as silicon and silicon nitride (*58*).

Because TFLN combines high confinement and high nonlinearity, it is particularly suitable for low-power continuous-wave and low-energy pulsed applications (*59*). For high-power or high-energy pulsed excitation, confinement or nonlinearity are less critical, and many other bulk materials such potassium titanyl phosphate (KTP) and β-BaB_2O_4 (BBO) are frequently used. However, one advantage of thin-film platforms, such as TFLN, over bulk solutions is the flexibility of engineering the waveguide geometry, and hence dispersion, which results in greater flexibility of phase and group velocity matching, enabling tailoring of components within integrated systems to access previously unachievable operating regimes.

Material properties and engineering concepts for bridging the EM spectrum

LN is one of several materials that has many attractive properties for generating and manipulating EM waves, including large nonlinear-optic, electro-optic, and piezo-electric coefficients, as illustrated in Fig. 2. LN can also support various engineering concepts to further enhance these effects, as shown in Fig. 3.

Nonlinear-optical effects

Nonlinear-optical effects can be used to generate new EM waves through second-order [$\chi^{(2)}$] and third-order [$\chi^{(3)}$] nonlinear-optical processes.

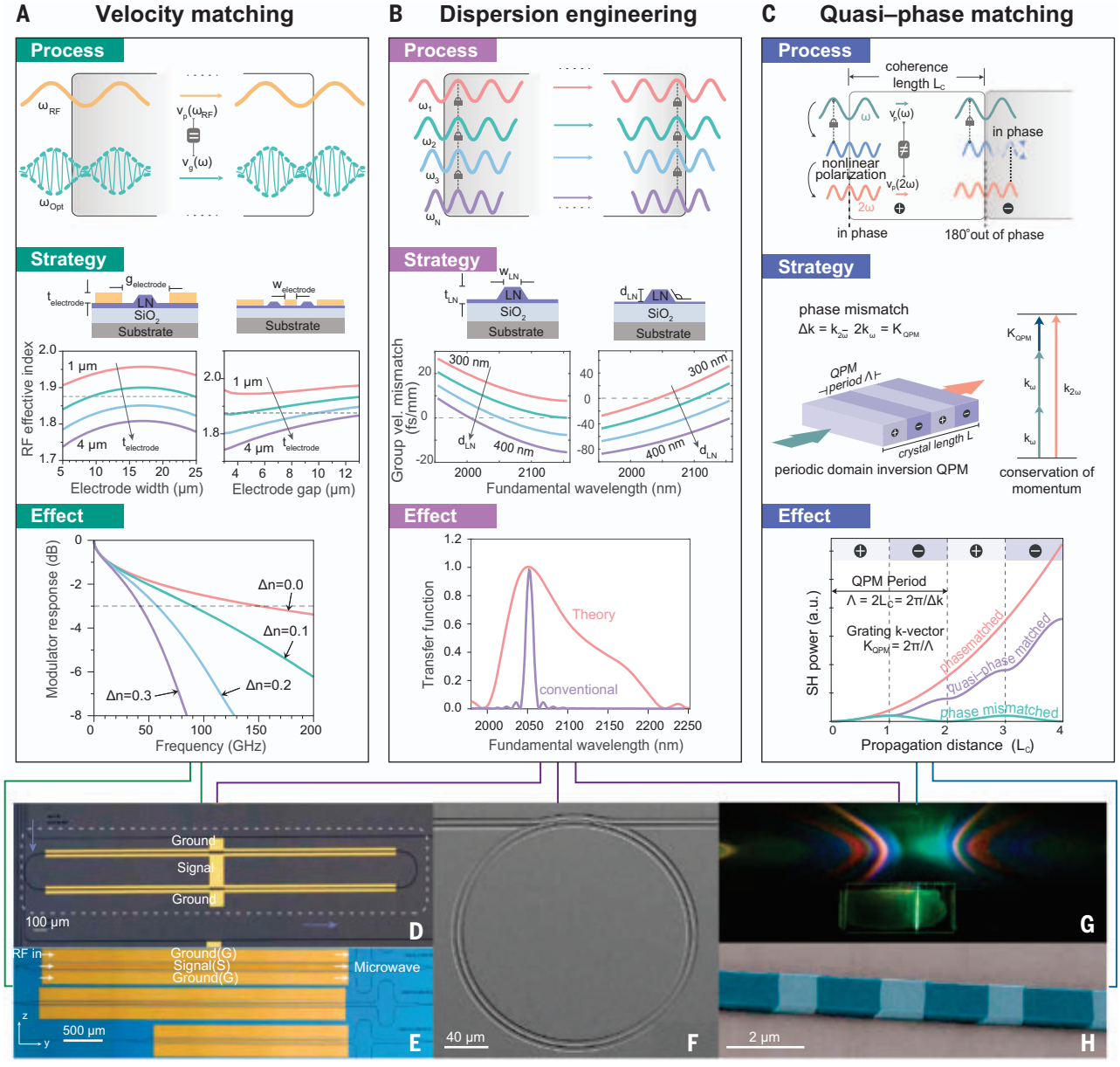


Fig. 3. Engineering concepts used in LN technology to enable efficient coupling for a wide spectral range. (**A** to **C**) Some of the most commonly used concepts in LN technology include the following: (i) Velocity matching (A) for broadband electro-optic interaction. Broadband electro-optic modulators can be realized by engineering the electrode dimensions to achieve velocity matching between the phase velocity of a single-frequency RF signal and the group velocity of the optical wave. (ii) Dispersion engineering (B) for broadband nonlinear optic interaction. Broadband nonlinear optical interaction can be realized by engineering the waveguide dimension to achieve low group velocity (vel.) mismatch. (iii) QPM (C) for efficient second-order nonlinear interaction. Phase mismatch between the interaction waves can be compensated to achieve efficient nonlinear optical interaction by inverting LN's spontaneous polarization periodically. a.u., arbitrary units. (**D**) EO comb generators that leverage velocity matching and dispersion engineering. [Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer *Nature* (*53*), 2019]. (**E**) High-speed EO modulators that leverage velocity matching. [Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer *Nature* (*49*), 2019]. (**F**) A Kerr microcomb generator that leverages dispersion engineering. [Reprinted with permission from (*52*) © The Optical Society]. (**G**) SHG waveguides that leverage dispersion engineering and QPM. [Reprinted with permission from (*59*) © The Optical Society]. (**H**) An SHG waveguide that leverages QPM. [Reprinted with permission from (*144*) © The Optical Society].

Below the Curie temperature of LN (1150°C), its crystal structure is noncentrosymmetric, giving rise to a large second-order nonlinearity. Meaning that an EM wave polarized along one crystal axis can cause a phase shift in another EM field with a certain polarization. The full set of such interactions is called the nonlinear tensor. For LN the strongest component is where both EM waves are polarized along the axis of crystal asymmetry(termed d_{33}) and is 27 pm/V at 1064 nm. Common second-order nonlinear processes include SHG, sum-frequency generation (SFG), and difference-frequency generation (DFG) (Fig. 2A). Third-order nonlinear optical processes in LN use its nonlinear refractive index (n_2) of 1.8×10^{-19} m^2/W (*60*), which is similar in strength to that of Si_3N_4 (2.5×10^{-19} m^2/W at 1.55 μm), enabling efficient four-wave mixing processes that are suitable for applications such as optical frequency comb (Fig. 2B).

Linear electro-optic effect

The linear electro-optic effect (or Pockels effect) changes the refractive index of LN proportional to an applied electric field, which can be used to modulate EM waves that pass through the crystal, generating new frequencies around the injected wave spaced by a single or multiple of the electrical modulation frequency (Fig. 2C). LN's largest electro-optic coefficient is crystal-axis dependent (closely related to the nonlinear tensor) with the largest component being on the order of 30 pm/V and typically requires interaction lengths of several millimeters up to a few centimeters to achieve the desired phase shifts with reasonably low voltages. The effect has been used to modulate and manipulate EM waves with electrical signals from static or very low frequency to hundreds of gigahertz. A key advantage of LN is that the electro-optic effect only changes the phase of light with no change to absorption. This is in contrast to silicon and other semiconductors where it is difficult to achieve independent phase and amplitude control.

Photo-elastic effect

The photo-elastic effect changes the refractive index of LN as a function of strain, which can be used for modulating and frequency-shifting EM waves by interacting with periodically sheared or compressed LN caused by the excitation of acoustic waves via the piezoelectric effect (Fig. 2D). Typically, operation frequencies of acousto-optical devices in LN are in the megahertz to single-digit gigahertz range, mainly because of practical considerations such as the size of the crystal and manufacturability of the electrodes.

Advantages of LN

A more detailed description of these effects and their use in LN can be found in several other review papers (*60*–*65*). Important to note is that LN does not necessarily exhibit the strongest material effects when compared with other materials. Indeed, there are many other materials such as KTP, BBO, GaAs, and InP that have attractive material properties. However, LN distinguishes itself by its maturity, stability, commercial availability, wide transparency range, and engineerability with respect to the coupling between EM frequencies within a wide spectral range, making their generation more efficient and tailorable.

Velocity matching

Velocity matching is used for one of the most common optical components in LN, namely broadband electro-optic traveling-wave modulators (Fig. 3A). For efficient electro-optic interaction of the RF wave with the optical wave, the phase velocity of the single-frequency RF wave and the group velocity of the optical wave should ideally be identical. This can be achieved by engineering waveguide and electrode dimensions as well as material dispersion. In the weakly confined LN waveguide platform, the relatively high dielectric constant ($\xi_{11,22}$ = 44, ξ_{33} = 27.9) of LN at microwave frequencies results in a velocity mismatch, which can be addressed by using thick electrodes (a few micrometers) to increase the phase velocity of the RF wave. The lower dielectric constant of SiO_2, which is a common buffer layer for tightly confining waveguides, increases the phase velocity of the RF wave, overcoming the need for thick electrodes for velocity matching.

Dispersion engineering

Dispersion engineering uses the material stack and waveguide dimensions as degrees of freedom to engineer the modal dispersion. Waveguide dispersion is determined by the wavelength-dependent field distribution in the core and cladding of the waveguide (Fig. 3B). This waveguide (or geometric) dispersion can become an important factor in the overall dispersion, otherwise solely determined by the material properties, by providing control over the group-velocity mismatch and group-velocity dispersion (*59*, *66*). Within the LN platforms mentioned above, meaningful dispersion engineering is usually only available in TFLN because of the large index contrast between core and cladding.

QPM

QPM is a technique that compensates for the phase-velocity (i.e., momentum) mismatch of different waves. This can be achieved by either periodically inverting the spontaneous polarization of the crystal when the phase mismatch reaches 180° [after a distance called the coherence length L_C (*17*)] (Fig. 3C) or, in guided-wave platforms, by periodically perturbing the magnitude of the nonlinear coupling through modifications of the waveguide dimensions (*67*). In the small signal conversion regime, this results in a unidirectional energy flow over the entire propagation length. The periodic reversal of the spontaneous polarization can be achieved by using well-developed optical, thermal, and electrical domain-engineering methods (*68*), among which the electric-field poling method is the most widely adopted one and has been applied to all three LN platforms to generate periodically and aperiodically poled LN crystals with periods reaching submicron dimensions (*69*). Because of the Fourier-transform relationship between the QPM grating and the device's transfer function (*70*–*72*), a nontrivial frequency response can also be readily engineered (*73*). It is important to note that electric-field poling and QPM engineerability are not available in most other optical material platforms, which is one of the main reasons for the wide adoption of LN for nonlinear optical applications.

LN photonics for unlocking the EM spectrum

LN's combination of material properties and engineerability provides the means to generate EM frequencies over a range that covers nearly five orders of magnitude (*53*, *74*–*77*), spanning from UV light to microwaves. Figure 4 illustrates the breadth of frequencies that have been experimentally generated and manipulated using LN as a linear and nonlinear frequency-mixing platform.

Visible and UV light

Visible and UV light (400 to 900 THz) experiences very low material losses in LN owing to the 3.93-eV-wide bandgap (~950 THz) (*78*) of LN. Light in this spectral range is required for applications such as virtual reality (*79*) and probing atomic transitions for optical clocks or magnetic field sensors (*80*), as well as for molecules and cells for bioimaging (*6*). Visible frequencies can be generated in LN by making use of the material's second- and third-order optical nonlinearities in combination with well-developed near-infrared light sources. For example, the second-order optical nonlinearity has enabled the generation of blue (*17*, *81*–*83*), green (*82*, *84*), yellow (*82*), orange (*82*, *85*), and red (*82*) frequencies by using either SHG, which increases the EM frequency by one octave; SFG; or a combination thereof. The highest EM frequencies that can be generated by such nonlinear optical processes are only limited by the UV absorption edge of LN (~950 THz, or 315 nm) and can reach up to 800 to 900 THz (*77*, *82*) in the near UV spectral range.

Supercontinuum generation (SCG) provides another means to generate visible light in LN by generating EM frequencies over a very broad spectral bandwidth that possibly covers the entire visible spectrum and can reach all the way to 850 THz (*59*, *76*, *86*). SCG typically uses the third-order nonlinearity in dispersion-engineered waveguides. However, in LN, the second-order optical nonlinearity can further help to push the generated spectrum toward shorter wavelengths through SHG and SFG (*76*). Furthermore, the cascading of two second-order processes can result in a large effective third-order nonlinearity, which exceeds the material's third-order nonlinearity and whose sign can be controlled through the choice of the sign of the phase mismatch. (*87*, *88*).

Near-infrared frequencies

Near-infrared frequencies (150 to 400 THz) are of particular interest because they are low enough for Rayleigh scattering to be minimal but high enough such that molecular absorption can be avoided in specific windows, which enables low loss transmission through optical fibers and on photonic integrated circuits (PICs). This makes them particularly attractive for applications such as optical communications (*1*, *2*), microwave photonics (*89*), and quantum

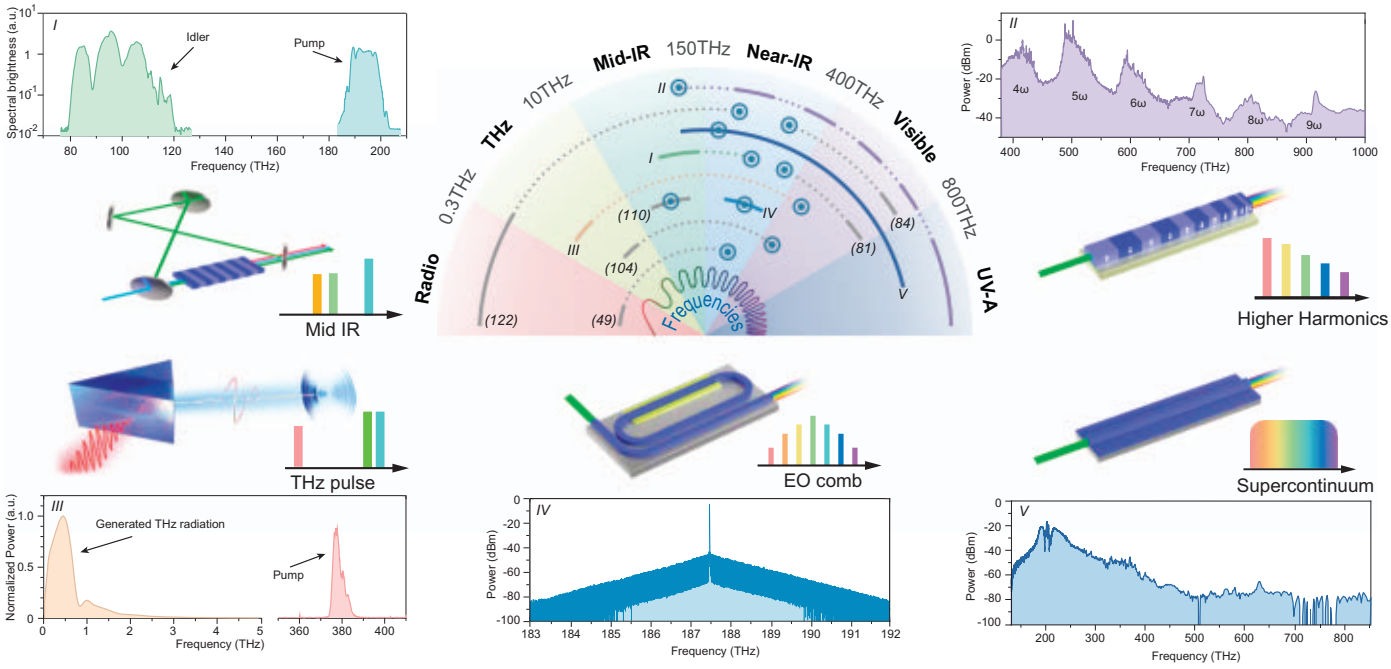


Fig. 4. LN spectral coverage. The semicircle illustrates the EM spectrum range and examples of generated EM frequencies when using LN for frequency mixing, with I to V representing OPO mid–infrared (IR) (I), higher harmonic (II), THz (III), EO comb (IV), and supercontinuum (V) generation [see (*49*, *81*, *84*, *104*, *110*, *122*) for examples]. The circled dots in the spectrum indicate the pump frequency. The illustrations and the detailed spectra surrounding the semicircle (from left to right, top to bottom) provide examples for (IR) (I) (*74*), (II) (*77*), (III) (*75*), (IV) (*53*), and (V) (*76*).

optics (*9*, *10*), as well as for free-space applications that operate outside the visible spectrum, such as light detection and ranging (LiDAR) and deep-space communication (*90*). Depending on the application, near-infrared frequencies can be generated in LN through a wide range of approaches, including Raman lasing and DFG, which are processes that generate one or two new EM frequencies, as well as Kerr microcombs, SCG, and electro-optic combs, which can generate tens to hundreds of new EM frequencies in the form of OFCs. For example, the generation of photons at 185 THz was demonstrated using Raman lasing in a high–Q-factor TFLN resonator (*91*) pumped at 192 THz. A popular approach for the generation of widely tunable EM frequencies is the use of an optical parametric oscillator (OPO). Here, two EM waves are generated, the signal and idler, which can be tuned by adjusting the phase-matching condition to enable the generation of frequencies between 200 and 268 THz (signal) and 110 and 178 THz (idler) (*92*). Such OPOs, operating with either continuous-wave or pulses of light, are commercially available and offer wide spectral tunability, which is difficult to achieve by other means. Squeezed light sources have also been demonstrated in the near-infrared spectral range as fundamental building blocks for quantum sensing and computing using degenerate parametric down conversion (*93*–*95*).

Applications that require the generation of many closely spaced EM frequencies can use OFCs. These combs have been demonstrated at near-infrared wavelengths using high–Q-factor, dispersion-engineered Kerr micro resonators in TFLN (*51*, *52*, *96*), which use the third-order optical nonlinearity of LN or cascaded second-order optical nonlinearities (*97*). Optical pump frequencies in the center of the C-band (~192 THz) are often used because of the availability of inexpensive light sources, such as narrow-linewidth and mode-locked lasers and high-power fiber amplifiers. An alternative way to generate near-infrared wavelengths is to use SCG (similar to the process outlined for visible-frequency generation), which has been used for generating OFCs that cover the full near-infrared frequency range and even reach into the visible (600 THz) and mid-infrared (100 THz) regions (*59*, *98*). Electro-optic combs provide another powerful tool to generate OFCs—hundreds of comb lines from 183 to 192 THz spaced by 25 GHz have been demonstrated (*53*, *99*–*101*).

Mid-infrared frequencies

Mid-infrared frequencies (10 to 150 THz) experience low losses in LN for frequencies above ~55 THz, when the phonon absorption starts to take place (*102*). Mid-infrared frequencies are attractive because they can be used to excite vibrational states of molecules (*103*) and hence are useful sources for spectroscopic sensors in applications such as air-quality monitoring in cities or for process monitoring of chemical plants and emissions from pipelines. Such frequencies have historically been generated in OPOs by using bulk periodically poled LN (PPLN) crystals. Use of this process has enabled the generation of EM waves from near-infrared frequencies down to 56 THz in the mid-infrared region (*104*–*106*), which is at the edge of the transparency window of LN. More recently, mid-infrared frequencies generated by DFG have been demonstrated in weakly confining (*107*, *108*) as well as TFLN waveguides, with continuous wave conversion efficiencies improving by one to two orders of magnitude. TFLN waveguides can also be used to efficiently translate near-infrared frequency combs to the mid-infrared region owing to the large conversion bandwidth afforded by dispersion engineering (*109*).

Kerr microcombs (*110*) and SCG (*98*, *111*) have both been used to generate OFCs in the mid-infrared region, with the former relying on high–Q-factor TFLN ring resonators to generate an OFC from 139 to 162 THz (*96*). The frequency combs generated by SCG have a wider spectral width and can reach all the way to 60 THz (*111*), a value limited by the phonon absorption of LN. In the TFLN platform, the material absorption of the silica cladding layer underneath the LN thin film can also limit the generated frequencies, which has recently been overcome by using sapphire as a low-index cladding material (*109*) or by undercutting the waveguide region (*112*).

THz frequencies

THz radiation (0.3 to 10 THz) can penetrate paper, plastics, and fabric and is therefore

attractive for standoff sensing and security imaging, such as at airports. LN crystals are attractive for the generation of narrow-band, high-intensity THz frequencies by optical rectification, which is a second-order nonlinear-optic process and can also be described by intrapulse DFG. In this process, the spectral frequency components of ultrashort optical pump pulses interact with each other and generate new EM-frequency components. This process can generate frequencies up to the difference between the blue and red edges of the pulse spectrum, so that a 10-fs pulse centered at 800 nm, for example, could generate frequencies up to ~44 THz (6.8 μm).

The pump pulses that have been used for optical rectification in bulk LN crystals are in the near-infrared region owing to the availability of high-power, ultrashort optical light sources, which commonly generate pulsed THz radiation with frequencies ranging from 0.2 to 4 THz (*75*, *113*–*116*), although higher frequency generation is possible, as indicated above. Pulse energies up to 1.4 mJ at 0.4 THz have been demonstrated by operating bulk PPLN crystals at cryogenic temperatures and using chirped laser pulses (*75*), with conversion efficiencies reaching up to 0.9% (*117*). The generation of continuous-wave THz radiation, tunable from 1.34 to 1.70 THz and 3.06 to 3.59 THz, has also been demonstrated using cascaded optical parametric processes in a singly resonant OPO using a bulk PPLN crystal for phase matching (*104*).

Microwave frequencies

Microwave frequencies (0.3 to 300 GHz) are used in applications such as 5G and 6G communication, radar, and radio astronomy. In recent years, "RF photonics" has become a widely used term to describe applications such as the generation of ultrastable microwave sources or the low-loss remoting of source and transmitter, which uses light to carry modulated signals to an antenna and subsequent conversion to microwaves. In the nonlinear optic context, such microwave frequencies can be generated directly through a DFG process in a whispering gallery mode resonator for the optical and microwave mode (*118*) or by using TFLN-superconductor hybrid electro-optic systems (*119*).

However, in most RF photonics applications, LN is used to translate microwave frequencies onto an optical carrier, which can then be transmitted and manipulated in the optical domain and subsequently generate microwave frequencies using a photodetector with an appropriate bandwidth. The main mechanisms of conversion from microwave to optical frequency rely on the electro-optic and acousto-optic effect in LN (see Fig. 2, B and C). Electro-optic modulators have been demonstrated in bulk crystals weakly confining waveguides (*38*), and strongly confining TFLN waveguides (*49*), with the latter being particularly attractive because of the attainable high modulation speeds and low drive voltages. The bandwidth of these electro-optic modulators can reach over 100 GHz (*49*, *64*, *120*, *121*) and even approach THz levels (*122*). Acousto-optic devices have also been demonstrated in bulk crystals, weakly confining waveguides (*123*) and strongly confining TFLN waveguides (*124*, *125*), with interaction frequencies typically ranging from MHz to GHz levels. Such acousto-optic devices are traditionally used to induce small frequency shifts, as beam deflectors, as well as moderately fast switches and power regulators (*123*–*126*). More recently, they have enabled quantum transduction for superconducting qubits (*119*, *127*), which is attributable to the small acoustic mode volume in nanophotonic resonators fabricated in TFLN.

Future of LN technology

LN, in all three modalities (bulk LN crystals, weakly confining, and strongly confining LN waveguides), is widely used for nonlinear optic, acoustic, and electro-optic processes to generate and manipulate EM frequencies over a wide spectral range. Since LN's inception, the material and manufacturing processes have matured, resulting in mostly discrete components that perform well-defined functions. In the future, bulk LN crystal components will remain important for EM-frequency generation across the spectrum, particularly for applications that require high optical powers, such as in high-power OPOs, free-space acousto-optic and electro-optic modulators, and Q-switches in laser cavities. However, for the LN waveguide platforms, we foresee a rapid acceleration of developments across two dimensions: (i) complexity and (ii) spectral breadth (Fig. 5). Complexity will transition from millimeter-scale single components on chip to micrometer-scale nanophotonic circuits followed by complex multilayer networks in which diverse materials are heterogeneously integrated with LN and co-packaged with electronic circuits. The spectral breadth of these devices will transition from operating primarily at near-infrared frequencies to generating and manipulating EM frequencies from visible to microwave frequencies on demand.

Near-term

In the near-term (next 5 years), bulk crystals and weakly confining waveguides will remain essential platforms for the generation of EM frequencies, particularly in the near-infrared and visible frequency region because they offer mature fabrication and packaging processes. The high-power handling capability and the low interface loss of such commercially available products make them particularly attractive for serving as individual devices in system demonstrations, for example, as part of self-referencing systems for optical clocks, in which the LN components can be connected to other sophisticated photonic infrastructure. Furthermore, the low-loss interfacing of weakly confined waveguides is important for quantum optics applications, for example, for the generation of photon pairs by spontaneous parametric down-conversion as well as for frequency translation between quantum nodes and long-haul fiber networks, in addition to heterogeneous intranode conversion. Although LN is a well-studied material, more work is needed to investigate the optimization of light-induced absorption changes (photochromic effect), particularly for operations at short wavelengths (*128*).

The wider adoption of tightly confining LN waveguides is, at present, mainly hampered by two shortfalls: (i) the poorly understood uniformity, repeatability, and reliability of the commercially available starting material and (ii) the lack of low-loss interfaces to standard optical fibers. Although there has been some progress in solving the interface problem for submicron waveguides in LN (*129*, *130*), they still require specialized fibers and high-precision multistep e-beam lithography and etching, which may not be practical or economical for large-scale, fully packaged solutions.

A high-volume application, such as optical short-distance data communication, will provide the commercial means to address both the shortfalls and maturation of TFLN technologies. It will also drive the reduction of TFLN wafer and manufacturing costs as well as increase diversity in TFLN suppliers. This high-volume communication application is motivated by the ever-increasing need for data center and cloud infrastructure, which requires increased speed, reduced power consumption, and lower cost for next-generation communication systems. Indeed, TFLN is an ideal candidate because of its low drive voltage, compatibility with CMOS electronics (*49*), and high bandwidth. Long-haul telecom applications may also become commercial drivers to mature the TFLN platform because the major challenge for this application is performance, where TFLN's demonstrated data (symbol) rate of 120 gigabaud and beyond is highly attractive (*49*). These devices can be drop-in replacements for existing guided-wave LN electro-optic modulators, allowing for low threshold commercial adoption while offering performance advantages. For LiDAR applications, the recent breakthrough of Pockels cavity lasers (*131*) by integrating III-V semiconductor gain sections with TFLN-based external cavities offers unrivaled frequency modulation speed and reconfigurability to integrated lasers; injection locking of a laser diode into an external LN ring modulator can also achieve fast frequency modulation while the wavelength tuning range is limited (*132*).

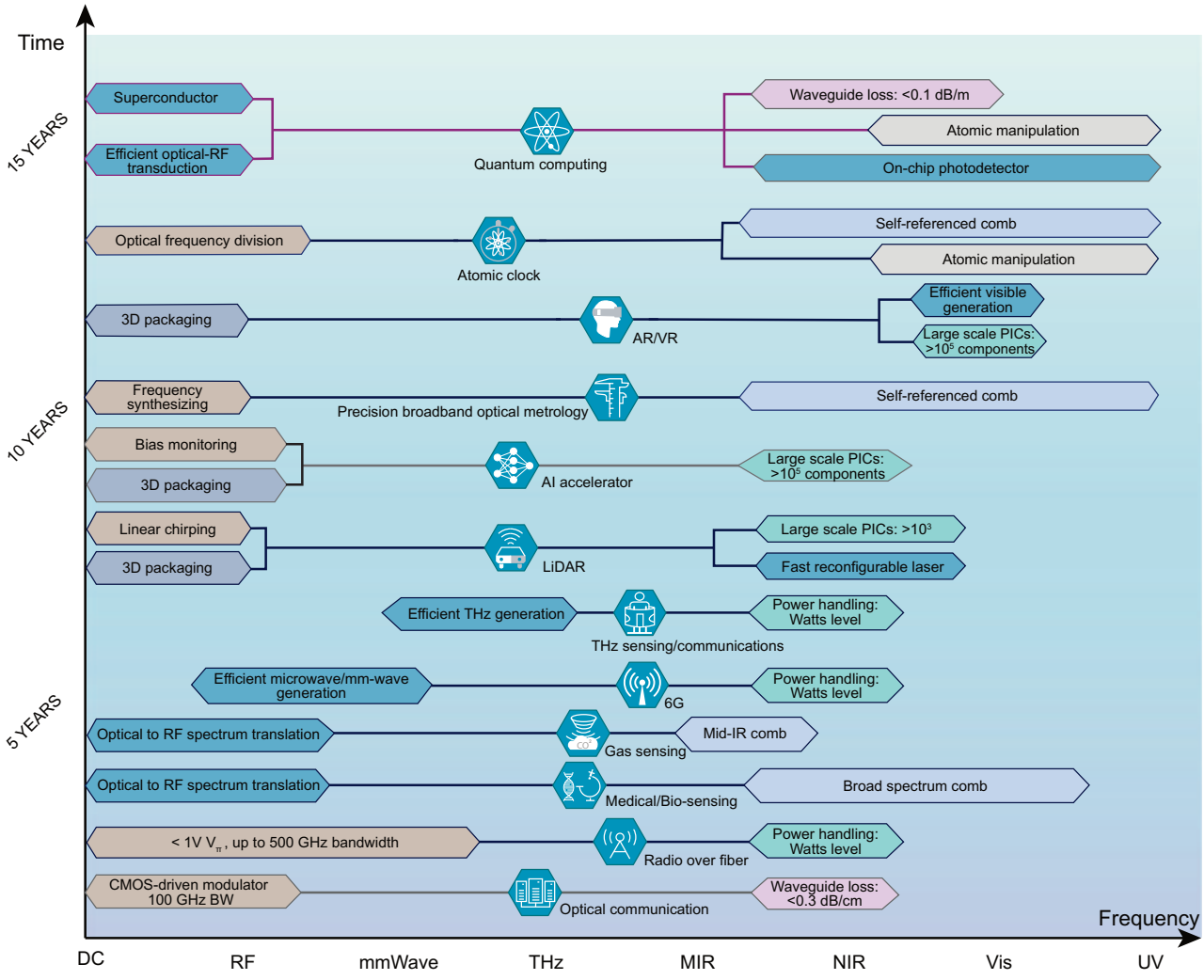


Fig. 5. Outlook of LN photonics and its applications. The applications are illustrated by the symbols in the center, which require the functionalities on either end of the spectrum. The continued development of quality engineering and production scalability drives the performance of photonic devices and systems, both of which jointly enable new applications. Emerging applications stimulate the improvement of material processing and high-volume device production. DC, direct current; MIR, mid-infrared; NIR, near-infrared; Vis, visible.

An alternative approach is the adiabatic frequency conversion using LN-based resonators (*133*). These approaches can potentially provide a chip-based, low-cost solution to frequency-modulated continuous-wave LiDAR. Furthermore, phase modulation with minimal intensity modulation in LN can be used to tune the emission angle of the light beam, without sensitivity degradation during the scanning. Addressing these near-term industrial applications with relatively simple, high-yield designs will help to mature the platform and achieve the statistical understanding of issues such as device yield, uniformity, repeatability, and reliability. They will also provide a better understanding of photorefractive and charging effects (e.g., bias drift in electro-optical modulators) in TFLN, which can have detrimental impacts on the performance and stability of LN PICs, particularly when their scale becomes large. We anticipate that in the next 5 years, substantial investments will be made in this space, which will build the foundation of the wide-scale adoption of tightly confining LN waveguide circuits as commercial devices.

At the same time, we foresee rapid progress in the development of TFLN at the device level for high-efficiency frequency mixing applications, such as miniaturized OPOs and optical phase arrays (OPAs) using integrated PPLN-based waveguides or microcavities, where earlier commercial adoption is feasible. For these applications, several devices can be manufactured on the same chip with parameter sweeps to account for uniformity and fabrication-tolerance issues, such that the best devices can be isolated and packaged (i.e., "hand-picked"). These devices will be of high value because they can provide ultrabroadband, ultrahigh parametric gains that reach an octave using microwatt average powers or femtojoule-level pump energies (*66*, *134*). The same approach can be used to develop widely tunable, on-chip oscillators for challenging spectral regimes, such as the visible and mid-infrared spectral ranges for addressing atomic and molecular systems used for metrology and spectroscopy, or efficient mm-wave and THz generation for wireless communication and sensing. Along this track of integrated OPOs and OPAs, squeezed-state generation for continuous-variable quantum computing, quantum random number generation, and quadratic dissipative soliton formation operating in a strongly nonlinear regime can be implemented.

One essential trend we expect to happen in the next 5 years is the development of heterogeneous integration on the TFLN platform. Based on wafer-bonding, LN waveguides will

be interfaced with different materials at the wafer scale, among which the integration of III-V semiconductor materials as laser sources, amplifiers, and detectors will be crucial. Another trend that we expect will be developed is the co-packaging of photonic and electronic circuits, which is crucial for the control and monitoring of large-scale LN circuits. Although these initial efforts will be mainly at the research level, they will pave the way for future wafer-scale heterogeneous platforms with full control and monitoring and communication functionality.

Mid-term

In the mid-term (5 to 10 years), we anticipate that bulk crystals and weakly confining waveguides will continue to be used as individual components for low-volume, price-sensitive visible to mid-infrared frequency applications. Emerging applications such as 6G or THz sensing will benefit from the high power-handling capability of these devices for efficient microwave and THz generation. However, the relatively high material losses at the THz spectral region (*135*, *136*) can be detrimental for device performance and require the generation of the THz radiation near the LN-air interface (*75*, *137*).

In parallel, we predict that improved low-loss optical interfaces to submicron waveguides and fabrication processes that are developed for high-volume data communication applications will result in an increased uptake of tightly confining TFLN waveguide circuits for near-infrared frequency applications. One of the challenges to be overcome in the mid-term is the reproducibility of nanophotonic components. For example, networks of virtually identical degenerate OPOs would be highly valuable for wavelength agile information processing. However, the high sensitivity of the nonlinear function of OPOs to the precise dispersion throughout the device imposes stringent requirements on the fabrication tolerance (waveguide widths and etch depth), wafer uniformity (film thickness, both within a single wafer as well from wafer-to-wafer) and also defect control. Active tuning and post-processing might hold solutions to these challenges. Hence, similar to silicon photonic circuits, one could rely on thermal heaters for tuning, which is not attractive for energy efficiency reasons, or the integration of advanced refractive index tunable materials (*138*). Where available, noncritical designs should be implemented to increase dimensional tolerances (*139*, *140*). This may be achievable when expanding the design space by considering sophisticated cladding structures, such as multilayer stacks.

Along with the improved fabrication processes and reproducibility of nanophotonic components, we anticipate that data communication applications will push forward the co-packaging of photonic and electronic circuits and the heterogeneous integration of active materials for near-infrared light sources and detectors (*57*), thereby transferring the fabrication processes that have been developed for silicon photonic circuits (*141*) to the TFLN platform. The integration of lasers and detectors will also be highly appealing to numerous applications that demand large-scale, fully integrated PICs, such as analog microwave photonics, LiDAR, and artificial intelligence (AI). Although at this stage the scalability of OPAs or optical neural networks (ONNs) might not be as vast as that available in silicon photonics, these applications will benefit immensely from the TFLN technology because of the substantially reduced drive voltage, increased bandwidth, and low insertion loss. Additionally, it is plausible to anticipate that complex and high-performance analog systems, including 6G mm-wave systems, RF spectrum analyzers, and photonic analog-to-digital converters, will be realized on TFLN. Importantly, these applications can benefit from the reduced cost by using the matured, near-infrared components to make the dream of the wide adoption of photonic-based, wideband-RF systems a reality.

Additionally, we envision a more complex system that operates from microwave wavelengths down to visible wavelengths and is enabled by further integration of electro-optic, acousto-optic, and all-optical signal-processing components. One of the foreseeable areas where LN can make a revolutionary impact is in OFC technology. As a material with both second-order and Kerr nonlinearity along with high-bandwidth modulators, LN is extremely suitable for realizing integrated self-referenced OFCs on a single nanophotonic platform with laser sources, amplifiers, and high-speed photodetectors that are heterogeneously integrated, where ultraefficient nonlinear broadening and SHG are readily accessible functions. In addition, the realization of cascading a series of low–switching voltage, high-bandwidth electro-optic modulators would promise, among many other applications, frequency-agile, electro-optic frequency combs with a considerable versatility in center wavelengths and repetition rates as well as direct synthesis of femtosecond-class pulse sources without mode locking (*142*). Such sources naturally operate at repetition rates that are compatible with on-chip resonators and thus open a wide range of opportunities in highly efficient nonlinear frequency conversion using pulsed pumping schemes. Other mid-term applications might involve dynamic beamforming based on electro- or acousto-optic devices, which is essential for LiDAR, augmented and virtual reality displays, and trapped-ion quantum computing systems. The major challenge in the mid-term is to establish innovative system architectures that can use different light-matter interactions (between microwave, mechanical, and optical subcomponents) without paying a substantial price in performance and manufacturing complexity.

Long-term

In the long-term (10 years and beyond), TFLN will be based on a large-scale (beyond 200-mm diameter wafer) foundry process with diverse heterogeneously integrated materials and co-packaged electronic circuits. Such a platform will be an excellent choice for scaling up optical networking schemes in which large arrays of classical or quantum light sources or processing units need to be coupled, often in a programmable way. This will enable fundamentally innovative applications such as fully integrated LiDAR and ONNs, quantum computing, fully integrated frequency synthesizers, massive RF signal processing networks, and advanced sensors. For example, photonic AI accelerators require an array of low–energy cost electro-optic modulators and nonlinear photonic activation components. In addition to neural networks, other photonic architectures for optical computing such as coherent Ising machines also demand a large number of photonic-based artificial spins connected by a reconfigurable all-to-all coupling matrix. Regardless of whether a time or spatially multiplexed approach is used, TFLN devices are critical to realizing photonic spins based on degenerate OPOs as well as spin-spin coupling based on delay lines with amplitude and phase modulators or a Mach-Zehnder interferometer mesh. Topological studies in synthetic dimensions could also be interesting in such coupled resonator networks.

A grand challenge in quantum engineering is to achieve extreme optical nonlinearity (ideally at the single-photon level) where both improvement of material processing and nanoengineering are required. We do see benefits from gradual improvements in optical nonlinearity using LN in future quantum PICs. This will benefit quantum photonic systems in which spatially multiplexed spontaneous parametric down-conversion sources with fast feedback optical switches are required for the implementation of near-deterministic single-photon sources or large scale continuous-variable computation in which high-quality squeezing states are required. Regarding quantum communication networks and all-optical signal processing, we envision that spectrotemporal shaping and quantum transduction techniques will have to be used to overcome the inhomogeneity of quantum emitters or to bridge the spectral difference between heterogeneous quantum systems. Furthermore, LN's fast modulation capability with low loss based on the electro-optic effect will be essential in almost all the quantum systems to increase processing speed and reduce system losses.

Fortunately, LN is an all-around high-performance optical material that is able to generate and manipulate EM frequencies on demand over a spectral range that covers nearly five orders of magnitude with a proven history of reliability and an ever-growing diverse portfolio of near- to long-term applications based on its nanophotonic platform. It is important to note that we anticipate that PICs that make use of TFLN will require heterogeneous integration to enable the integration of the attractive material properties of LN with efficient light sources and detectors. This may be realized by heterogeneously integrating such PIC elements on the TFLN waveguide platform (*57*) or by integrating thin films of LN onto other PIC material platforms (*143*). Independent of its modality (bulk crystal, weakly confined, or tightly confined waveguides) or, indeed, integration method, LN is in a great position to overcome the outlined challenges to become the material platform of choice for unlocking the EM spectrum and to continue to revolutionize optical science for years to come.

REFERENCES AND NOTES

1. P. Marin-Palomo *et al*., Microresonator-based solitons for massively parallel coherent optical communications. *Nature* **546**, 274–279 (2017). doi: 10.1038/nature22387; pmid: 28593968
2. B. Corcoran *et al*., Ultra-dense optical data transmission over standard fibre with a single chip source. *Nat. Commun.* **11**, 2568 (2020). doi: 10.1038/s41467-020-16265-x; pmid: 32444605
3. T. W. Hänsch, Nobel lecture: Passion for precision. *Rev. Mod. Phys.* **78**, 1297–1309 (2006). doi: 10.1103/RevModPhys.78.1297
4. J. L. Hall, Nobel lecture: Defining and measuring optical frequencies. *Rev. Mod. Phys.* **78**, 1279–1295 (2006). doi: 10.1103/RevModPhys.78.1279; pmid: 17086589
5. H. R. Telle *et al*., Carrier-envelope offset phase control: A novel concept for absolute optical frequency measurement and ultrashort pulse generation. *Appl. Phys. B* **69**, 327–332 (1999). doi: 10.1007/s003400050813
6. R. Prevedel, A. Diz-Muñoz, G. Ruocco, G. Antonacci, Brillouin microscopy: An emerging tool for mechanobiology. *Nat. Methods* **16**, 969–977 (2019). doi: 10.1038/s41592-019-0543-3; pmid: 31548707
7. J. G. Fujimoto, Optical coherence tomography for ultrahigh resolution in vivo imaging. *Nat. Biotechnol.* **21**, 1361–1367 (2003). doi: 10.1038/nbt892; pmid: 14595364
8. U. Morgner *et al*., Spectroscopic optical coherence tomography. *Opt. Lett.* **25**, 111–113 (2000). doi: 10.1364/OL.25.000111; pmid: 18059799
9. T. D. Ladd *et al*., Quantum computers. *Nature* **464**, 45–53 (2010). doi: 10.1038/nature08812; pmid: 20203602
10. S. Wehner, D. Elkouss, R. Hanson, Quantum internet: A vision for the road ahead. *Science* **362**, eaam9288 (2018). doi: 10.1126/science.aam9288; pmid: 30337383
11. M. Manzo, F. Laurell, V. Pasiskevicius, K. Gallo, "Lithium niobate: The silicon of photonics!" in *Nano-Optics for Enhancing Light-Matter Interactions on a Molecular Scale*, B. Di Bartolo, J. Collins, Eds. (NATO Science for Peace and Security Series B: Physics and Biophysics, Springer, 2013), pp. 421–422.
12. B. T. Matthias, J. P. Remeika, Ferroelectricity in the ilmenite structure. *Phys. Rev.* **76**, 1886–1887 (1949). doi: 10.1103/PhysRev.76.1886.2
13. G. E. Peterson, A. A. Ballman, P. V. Lenzo, P. M. Bridenbaugh, Electro-optic properties of LiNbO3. *Appl. Phys. Lett.* **5**, 62–64 (1964). doi: 10.1063/1.1754053
14. R. G. Smith, K. Nassau, M. F. Galvin, Efficient continuous optical second-harmonic generation. *Appl. Phys. Lett.* **7**, 256–258 (1965). doi: 10.1063/1.1754246
15. K. Nassau, H. J. Levinstein, G. M. Loiacono, Ferroelectric lithium niobate. 1. Growth, domain structure, dislocations and etching. *J. Phys. Chem. Solids* **27**, 983–988 (1966). doi: 10.1016/0022-3697(66)90070-9
16. K. Nassau, H. J. Levinstein, G. M. Loiacono, Ferroelectric lithium niobate. 2. Preparation of single domain crystals. *J. Phys. Chem. Solids* **27**, 989–996 (1966). doi: 10.1016/0022-3697(66)90071-0
17. M. Yamada, N. Nada, M. Saitoh, K. Watanabe, First-order quasi-phase matched LiNbO3 waveguide periodically poled by applying an external field for efficient blue second-harmonic generation. *Appl. Phys. Lett.* **62**, 435–436 (1993). doi: 10.1063/1.108925
18. L. Huang, N. A. F. Jaeger, Discussion of domain inversion in LiNbO3. *Appl. Phys. Lett.* **65**, 1763–1765 (1994). doi: 10.1063/1.112911
19. A. Ashkin *et al*., Optically-induced refractive index inhomogeneities in $LiNbO_3$ AND $LiTaO_3$. *Appl. Phys. Lett.* **9**, 72–74 (1966). doi: 10.1063/1.1754607
20. F. H. Mok, M. C. Tackitt, H. M. Stoll, Storage of 500 high-resolution holograms in a $LiNbO_3$ crystal. *Opt. Lett.* **16**, 605–607 (1991). doi: 10.1364/OL.16.000605; pmid: 19774013
21. C. R. Phillips, J. S. Pelc, M. M. Fejer, Continuous wave monolithic quasi-phase-matched optical parametric oscillator in periodically poled lithium niobate. *Opt. Lett.* **36**, 2973–2975 (2011). doi: 10.1364/OL.36.002973; pmid: 21808376
22. T. Beckmann *et al*., Highly tunable low-threshold optical parametric oscillation in radially poled whispering gallery resonators. *Phys. Rev. Lett.* **106**, 143903 (2011). doi: 10.1103/PhysRevLett.106.143903; pmid: 21561193
23. S. Schiller, R. L. Byer, Quadruply resonant optical parametric oscillation in a monolithic total-internal-reflection resonator. *J. Opt. Soc. Am. B* **10**, 1696 (1993). doi: 10.1364/JOSAB.10.001696
24. V. S. Ilchenko, A. B. Matsko, A. A. Savchenkov, L. Maleki, "High-efficiency microwave and millimeter-wave electro-optical modulation with whispering-gallery resonators" in *Laser Resonators and Beam Control V*, vol. 4629 of *SPIE Proceedings*, A. V. Kudryashov, Ed. (SPIE, 2002), pp. 158–163.
25. V. S. Ilchenko, A. A. Savchenkov, A. B. Matsko, L. Maleki, Nonlinear optics and crystalline whispering gallery mode cavities. *Phys. Rev. Lett.* **92**, 043903 (2004). doi: 10.1103/PhysRevLett.92.043903; pmid: 14995375
26. D. Wei *et al*., Experimental demonstration of a three-dimensional lithium niobate nonlinear photonic crystal. *Nat. Photonics* **12**, 596–600 (2018). doi: 10.1038/s41566-018-0240-2
27. L. E. Myers *et al*., Quasi-phase-matched 1.064-microm-pumped optical parametric oscillator in bulk periodically poled $LiNbO_3$. *Opt. Lett.* **20**, 52–54 (1995). doi: 10.1364/OL.20.000052; pmid: 19855794
28. U. Strößner *et al*., Single-frequency continuous-wave optical parametric oscillator system with an ultrawide tuning range of 550 to 2830 nm. *J. Opt. Soc. Am. B* **19**, 1419 (2002). doi: 10.1364/JOSAB.19.001419
29. I. P. Kaminow, J. R. Carruthers, Optical waveguiding layers in $LiNbO_3$ and $LiTaO_3$. *Appl. Phys. Lett.* **22**, 326–328 (1973). doi: 10.1063/1.1654657
30. R. V. Schmidt, I. P. Kaminow, Metal-diffused optical waveguides in $LiNbO_3$. *Appl. Phys. Lett.* **25**, 458–460 (1974). doi: 10.1063/1.1655547
31. J. L. Jackel, C. E. Rice, J. J. Veselka, Proton exchange for high-index waveguides in $LiNbO_3$. *Appl. Phys. Lett.* **41**, 607–608 (1982). doi: 10.1063/1.93615
32. L. Gui, B. Xu, T. C. Chong, Microstructure in lithium niobate by use of focused femtosecond laser pulses. *IEEE Photonics Technol. Lett.* **16**, 1337–1339 (2004). doi: 10.1109/LPT.2004.826112
33. G. Bava, I. Montrosset, W. Sohler, H. Suche, Numerical modeling of Ti:$LiNbO_3$ integrated optical parametric oscillators. *IEEE J. Quantum Electron.* **23**, 42–51 (1987). doi: 10.1109/JQE.1987.1073210
34. Y. N. Korkishko, V. A. Fedorov, Relationship between refractive indices and hydrogen concentration in proton exchanged $LiNbO_3$ waveguides. *J. Appl. Phys.* **82**, 1010–1017 (1997). doi: 10.1063/1.365864
35. J. Webjorn, F. Laurell, G. Arvidsson, Fabrication of periodically domain-inverted channel waveguides in lithium niobate for second harmonic generation. *J. Lightwave Technol.* **7**, 1597–1600 (1989). doi: 10.1109/50.39103
36. E. J. Lim, M. M. Fejer, R. L. Byer, Second-harmonic generation of green light in periodically poled planar lithium niobate waveguide. *Electron. Lett.* **25**, 174 (1989). doi: 10.1049/el:19890127
37. R. Brinkmann, W. Sohler, H. Suche, Continuous-wave erbium-diffused $LiNbO_3$ waveguide laser. *Electron. Lett.* **27**, 415 (1991). doi: 10.1049/el:19910263
38. E. L. Wooten *et al*., Review of lithium niobate modulators for fiber-optic communications systems. *IEEE J. Sel. Top. Quantum Electron.* **6**, 69–82 (2000). doi: 10.1109/2944.826874
39. C. Langrock *et al*., Highly efficient single-photon detection at communication wavelengths by use of upconversion in reverse-proton-exchanged periodically poled $LiNbO_3$ waveguides. *Opt. Lett.* **30**, 1725–1727 (2005). doi: 10.1364/OL.30.001725; pmid: 16075551
40. H. Jin *et al*., On-chip generation and manipulation of entangled photons based on reconfigurable lithium-niobate waveguide circuits. *Phys. Rev. Lett.* **113**, 103601 (2014). doi: 10.1103/PhysRevLett.113.103601; pmid: 25238358
41. J. Gil-Lopez *et al*., Improved non-linear devices for quantum applications. *New J. Phys.* **23**, 063082 (2021). doi: 10.1088/1367-2630/ac09fd
42. C. Langrock, S. Kumar, J. E. McGeehan, A. E. Willner, M. M. Fejer, All-optical signal processing using /spl chi//sup (2)/ nonlinearities in guided-wave devices. *J. Lightwave Technol.* **24**, 2579–2592 (2006). doi: 10.1109/JLT.2006.874605
43. P. Rabiei, P. Gunter, Optical and electro-optical properties of submicrometer lithium niobate slab waveguides prepared by crystal ion slicing and wafer bonding. *Appl. Phys. Lett.* **85**, 4603–4605 (2004). doi: 10.1063/1.1819527
44. H. Hu, R. Ricken, W. Sohler, Lithium niobate photonic wires. *Opt. Express* **17**, 24261–24268 (2009). doi: 10.1364/OE.17.024261; pmid: 20052137
45. L. Chang *et al*., Thin film wavelength converters for photonic integrated circuits. *Optica* **3**, 531 (2016). doi: 10.1364/OPTICA.3.000531
46. M. Zhang, C. Wang, R. Cheng, A. Shams-Ansari, M. Lončar, Monolithic ultra-high-Q lithium niobate microring resonator. *Optica* **4**, 1536 (2017). doi: 10.1364/OPTICA.4.001536
47. J. Lu *et al*., Periodically poled thin-film lithium niobate microring resonators with a second-harmonic generation efficiency of 250,000%/W. *Optica* **6**, 1455 (2019). doi: 10.1364/OPTICA.6.001455
48. R. Wolf *et al*., Quasi-phase-matched nonlinear optical frequency conversion in on-chip whispering galleries. *Optica* **5**, 872 (2018). doi: 10.1364/OPTICA.5.000872
49. C. Wang *et al*., Integrated lithium niobate electro-optic modulators operating at CMOS-compatible voltages. *Nature* **562**, 101–104 (2018). doi: 10.1038/s41586-018-0551-y; pmid: 30250251
50. K. Aoki *et al*., Single-drive X-cut thin-sheet $LiNbO_3$ optical modulator with chirp adjusted using asymmetric CPW electrode. *J. Lightwave Technol.* **24**, 2233–2237 (2006). doi: 10.1109/JLT.2006.872303
51. C. Wang *et al*., Monolithic lithium niobate photonic circuits for Kerr frequency comb generation and modulation. *Nat. Commun.* **10**, 978 (2019). doi: 10.1038/s41467-019-08969-6; pmid: 30816151
52. Y. He *et al*., Self-starting bi-chromatic $LiNbO_3$ soliton microcomb. *Optica* **6**, 1138 (2019). doi: 10.1364/OPTICA.6.001138
53. M. Zhang *et al*., Broadband electro-optic frequency comb generation in a lithium niobate microring resonator. *Nature* **568**, 373–377 (2019). doi: 10.1038/s41586-019-1008-7; pmid: 30858615
54. X. Liu *et al*., Tunable single-mode laser on thin film lithium niobate. *Opt. Lett.* **46**, 5505–5508 (2021). doi: 10.1364/OL.441167; pmid: 34724512
55. Y. Liu *et al*., On-chip erbium-doped lithium niobate microcavity laser. *Sci. China Phys. Mech. Astron.* **64**, 234262 (2021). doi: 10.1007/s11433-020-1625-9
56. K. Luke *et al*., Wafer-scale low-loss lithium niobate photonic integrated circuits. *Opt. Express* **28**, 24452–24458 (2020). doi: 10.1364/OE.401959; pmid: 32906986
57. C. Op de Beeck *et al*., III/V-on-lithium niobate amplifiers and lasers. *Optica* **8**, 1288 (2021).
58. T. Komljenovic *et al*., Photonic integrated circuits using heterogeneous integration on silicon. *Proc. IEEE Inst. Electr. Electron. Eng.* **106**, 2246–2257 (2018).
59. M. Jankowski *et al*., Ultrabroadband nonlinear optics in nanophotonic periodically poled lithium niobate waveguides. *Optica* **7**, 40 (2020). doi: 10.1364/OPTICA.7.000040

60. D. Zhu *et al*., Integrated photonics on thin-film lithium niobate. *Adv. Opt. Photonics* **13**, 242 (2021). doi: 10.1364/AOP.411024
61. L. Arizmendi, Photonic applications of lithium niobate crystals. *Phys. Status Solidi, A Appl. Res.* **201**, 253–283 (2004). doi: 10.1002/pssa.200303911
62. G. Chen *et al*., Advances in lithium niobate photonics: Development status and perspectives. *Adv. Photonics* **4**, (2022). doi: 10.1117/1.AP.4.3.034003
63. M. G. Vazimali, S. Fathpour, Applications of thin-film lithium niobate in nonlinear integrated photonics. *Adv. Photonics* **4**, (2022). doi: 10.1117/1.AP.4.3.034001
64. M. Zhang, C. Wang, P. Kharel, D. Zhu, M. Lončar, Integrated lithium niobate electro-optic modulators: When performance meets scalability. *Optica* **8**, 652 (2021). doi: 10.1364/OPTICA.415762
65. Y. Qi, Y. Li, Integrated lithium niobate photonics. *Nanophotonics* **9**, 1287–1320 (2020). doi: 10.1515/nanoph-2020-0013
66. L. Ledezma *et al*., Intense optical parametric amplification in dispersion-engineered nanophotonic lithium niobate waveguides. *Optica* **9**, 303 (2022). doi: 10.1364/OPTICA.442332
67. C. Wang *et al*., Second harmonic generation in nano-structured thin-film lithium niobate waveguides. *Opt. Express* **25**, 6963–6973 (2017). doi: 10.1364/OE.25.006963; pmid: 28381038
68. V. Y. Shur, A. R. Akhmatkhanov, I. S. Baturin, Micro- and nano-domain engineering in lithium niobate. *Appl. Phys. Rev.* **2**, 040604 (2015). doi: 10.1063/1.4928591
69. G. Rosenman, P. Urenski, A. Agronin, Y. Rosenwaks, M. Molotskii, Submicron ferroelectric domain structures tailored by high-voltage scanning probe microscopy. *Appl. Phys. Lett.* **82**, 103–105 (2003). doi: 10.1063/1.1534410
70. G. Imeshev *et al*., Engineerable femtosecond pulse shaping by second-harmonic generation with Fourier synthetic quasi-phase-matching gratings. *Opt. Lett.* **23**, 864–866 (1998). doi: 10.1364/OL.23.000864; pmid: 18087367
71. M. Asobe *et al*., Engineered quasi-phase matching device for unequally spaced multiple wavelength generation and its application to midinfrared gas sensing. *IEEE J. Quantum Electron.* **46**, 447–453 (2010). doi: 10.1109/JQE.2009.2032366
72. G. Imeshev *et al*., Ultrashort-pulse second-harmonic generation with longitudinally nonuniform quasi-phase-matching gratings: Pulse compression and shaping. *J. Opt. Soc. Am. B* **17**, 304 (2000). doi: 10.1364/JOSAB.17.000304
73. C. J. Xin *et al*., Spectrally separable photon-pair generation in dispersion engineered thin-film lithium niobate. *Opt. Lett.* **47**, 2830–2833 (2022). doi: 10.1364/OL.456873; pmid: 35648941
74. N. Leindecker, A. Marandi, R. L. Byer, K. L. Vodopyanov, Broadband degenerate OPO for mid-infrared frequency comb generation. *Opt. Express* **19**, 6296–6302 (2011). doi: 10.1364/OE.19.006296; pmid: 21451655
75. B. Zhang *et al*., 1.4-mJ high energy terahertz radiation from lithium niobates. *Laser Photonics Rev.* **15**, 2000295 (2021). doi: 10.1002/lpor.202000295
76. M. Yu, B. Desiatov, Y. Okawachi, A. L. Gaeta, M. Lončar, Coherent two-octave-spanning supercontinuum generation in lithium-niobate waveguides. *Opt. Lett.* **44**, 1222–1225 (2019). doi: 10.1364/OL.44.001222; pmid: 30821753
77. J. Rutledge *et al*., Broadband ultraviolet-visible frequency combs from cascaded high-harmonic generation in quasi-phase-matched waveguides. *J. Opt. Soc. Am. B* **38**, 2252 (2021). doi: 10.1364/JOSAB.427086
78. C. Nico, T. Monteiro, M. P. F. Graça, Niobium oxides and niobates physical properties: Review and prospects. *Prog. Mater. Sci.* **80**, 1–37 (2016). doi: 10.1016/j.pmatsci.2016.02.001
79. M. Hu *et al*., Virtual reality: A survey of enabling technologies and its applications in IoT. *J. Netw. Comput. Appl.* **178**, 102970 (2021). doi: 10.1016/j.jnca.2020.102970
80. N. Chauhan *et al*., Ultra-low loss visible light waveguides for integrated atomic, molecular, and quantum photonics. *Opt. Express* **30**, 6960–6969 (2022). doi: 10.1364/OE.448938; pmid: 35299469
81. A. Al Sayem *et al*., Efficient and tunable blue light generation using lithium niobate nonlinear photonics. *Appl. Phys. Lett.* **119**, 231104 (2021). doi: 10.1063/5.0071769
82. J. Capmany, Simultaneous generation of red, green, and blue continuous-wave laser radiation in Nd3 -doped aperiodically poled lithium niobate. *Appl. Phys. Lett.* **78**, 144–146 (2001). doi: 10.1063/1.1338495
83. L. Carletti *et al*., Steering and encoding the polarization of the second harmonic in the visible with a monolithic $LiNbO_3$ metasurface. *ACS Photonics* **8**, 731–737 (2021). doi: 10.1021/acsphotonics.1c00026; pmid: 33842671
84. G. D. Miller, R. G. Batchko, M. M. Fejer, R. L. Byer, "Visible quasi-phase-matched harmonic generation by electric-field-poled lithium niobate" in *Nonlinear Frequency Generation and Conversion*, vol. 2700 of *SPIE Proceedings*, M. C. Gupta, W. J. Kozlovsky, D. C. MacPherson, Eds. (SPIE, 1996).
85. S. Sinha, C. Langrock, M. J. F. Digonnet, M. M. Fejer, R. L. Byer, Efficient yellow-light generation by frequency doubling a narrow-linewidth 1150 nm ytterbium fiber oscillator. *Opt. Lett.* **31**, 347–349 (2006). doi: 10.1364/OL.31.000347; pmid: 16480204
86. M. Reig Escalé, F. Kaufmann, H. Jiang, D. Pohl, R. Grange, Generation of 280 THz-spanning near-ultraviolet light in lithium niobate-on-insulator waveguides with sub-100 pJ pulses. *APL Photonics* **5**, 121301 (2020). doi: 10.1063/5.0028776
87. R. Normandin, G. I. Stegeman, Nondegenerate four-wave mixing in integrated optics. *Opt. Lett.* **4**, 58 (1979). doi: 10.1364/OL.4.000058; pmid: 19684783
88. S. Liu *et al*., Effective four-wave mixing in the lithium niobate on insulator microdisk by cascading quadratic processes. *Opt. Lett.* **44**, 1456–1459 (2019). doi: 10.1364/OL.44.001456; pmid: 30874675
89. D. Marpaung, J. Yao, J. Capmany, Integrated microwave photonics. *Nat. Photonics* **13**, 80–90 (2019). doi: 10.1038/s41566-018-0310-5
90. J. Riemensberger *et al*., "Massively parallel coherent LiDAR using dissipative Kerr solitons" in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (Optica Publishing Group, 2020), paper SM2N.2.doi: 10.1364/CLEO_SI.2020.SM2N.2
91. M. Yu *et al*., Raman lasing and soliton mode-locking in lithium niobate microresonators. *Light Sci. Appl.* **9**, 9 (2020). doi: 10.1038/s41377-020-0246-7;pmid: 31969982
92. K. C. Burr, C. L. Tang, M. A. Arbore, M. M. Fejer, High-repetition-rate femtosecond optical parametric oscillator based on periodically poled lithium niobate. *Appl. Phys. Lett.* **70**, 3341–3343 (1997). doi: 10.1063/1.119164
93. P.-K. Chen, I. Briggs, S. Hou, L. Fan, Ultra-broadband quadrature squeezing with thin-film lithium niobate nanophotonics. *Opt. Lett.* **47**, 1506–1509 (2022). doi: 10.1364/OL.447695; pmid: 35290350
94. V. D. Vaidya *et al*., Broadband quadrature-squeezed vacuum and nonclassical photon number correlations from a nanophotonic device. *Sci. Adv.* **6**, eaba9186 (2020). doi: 10.1126/sciadv.aba9186; pmid: 32967824
95. T. Kashiwazaki *et al*., Continuous-wave 6-dB-squeezed light with 2.5-THz-bandwidth from single-mode PPLN waveguide. *APL Photonics* **5**, 036104 (2020). doi: 10.1063/1.5142437
96. Z. Gong, X. Liu, Y. Xu, H. X. Tang, Near-octave lithium niobate soliton microcomb. *Optica* **7**, 1275 (2020). doi: 10.1364/OPTICA.400994
97. R. Wolf, I. Breunig, H. Zappe, K. Buse, Cascaded second-order optical nonlinearities in on-chip micro rings. *Opt. Express* **25**, 29927–29933 (2017). doi: 10.1364/OE.25.029927; pmid: 29221028
98. C. R. Phillips *et al*., Supercontinuum generation in quasi-phase-matched $LiNbO_3$ waveguide pumped by a Tm-doped fiber laser system. *Opt. Lett.* **36**, 3912–3914 (2011). doi: 10.1364/OL.36.003912; pmid: 21964139
99. A. Rueda, F. Sedlmeir, M. Kumari, G. Leuchs, H. G. L. Schwefel, Resonant electro-optic frequency comb. *Nature* **568**, 378–381 (2019). doi: 10.1038/s41586-019-1110-x; pmid: 30996319
100. A. J. Metcalf, V. Torres-Company, D. E. Leaird, A. M. Weiner, High-power broadly tunable electrooptic frequency comb generator. *IEEE J. Sel. Top. Quantum Electron.* **19**, 231–236 (2013). doi: 10.1109/JSTQE.2013.2268384
101. M. Lončar *et al*., Integrated high-efficiency and broadband electro-optic frequency comb generators. *Nat. Photon.* **16**, 679–685 (2022). doi: 10.1109/JSTQE.2013.2268384
102. M. Leidinger *et al*., Comparative study on three highly sensitive absorption measurement techniques characterizing lithium niobate over its entire transparent spectral range. *Opt. Express* **23**, 21690–21705 (2015). doi: 10.1364/OE.23.021690; pmid: 26368148
103. T. Shimanouchi, *Tables of Molecular Vibrational Frequencies, Consolidated Volume I* (NSRDS-NBS 39, Secretary of Commerce, 1972).
104. J. Kiessling, R. Sowade, I. Breunig, K. Buse, V. Dierolf, Cascaded optical parametric oscillations generating tunable terahertz waves in periodically poled lithium niobate crystals. *Opt. Express* **17**, 87–91 (2009). doi: 10.1364/OE.17.000087; pmid: 19129876
105. I. Galli *et al*., Ti:sapphire laser intracavity difference-frequency generation of 30 mW cw radiation around 4.5 μm. *Opt. Lett.* **35**, 3616–3618 (2010). doi: 10.1364/OL.35.003616; pmid: 21042368
106. T. Beddard, M. Ebrahimzadeh, T. D. Reid, W. Sibbett, Five-optical-cycle pulse generation in the mid infrared from an optical parametric oscillator based on aperiodically poled lithium niobate. *Opt. Lett.* **25**, 1052–1054 (2000). doi: 10.1364/OL.25.001052; pmid: 18064270
107. D. Richter *et al*., High-power, tunable difference frequency generation source for absorption spectroscopy based on a ridge waveguide periodically poled lithium niobate crystal. *Opt. Express* **15**, 564–571 (2007). doi: 10.1364/OE.15.000564; pmid: 19532275
108. K.-D. F. Büchter, H. Herrmann, C. Langrock, M. M. Fejer, W. Sohler, All-optical Ti:PPLN wavelength conversion modules for free-space optical transmission links in the mid-infrared. *Opt. Lett.* **34**, 470–472 (2009). doi: 10.1364/OL.34.000470; pmid: 19373344
109. J. Mishra *et al*., Mid-infrared nonlinear optics in thin-film lithium niobate on sapphire. *OSA Nonlinear Optics* **2021**, NW3A.5 (2021). doi: 10.1364/NLO.2021.NW3A.5
110. Z. Gong *et al*., Soliton microcomb generation at 2 μm in z-cut lithium niobate microring resonators. *Opt. Lett.* **44**, 3182–3185 (2019). doi: 10.1364/OL.44.003182; pmid: 31199411
111. A. S. Kowligy *et al*., Mid-infrared frequency comb generation via cascaded quadratic nonlinearities in quasi-phase-matched waveguides. *Opt. Lett.* **43**, 1678–1681 (2018). doi: 10.1364/OL.43.001678; pmid: 29652338
112. X. Liu *et al*., Sub-terahertz bandwidth capactively-loaded thin-film lithium niobate electro-optic modulators based on an undercut structure. *Opt. Express* **29**, 41798 (2021). doi: 10.1364/OE.442091
113. Q. Tian *et al*., Efficient generation of a high-field terahertz pulse train in bulk lithium niobate crystals by optical rectification. *Opt. Express* **29**, 9624–9634 (2021). doi: 10.1364/OE.419709; pmid: 33820386
114. W. R. Huang *et al*., Highly efficient terahertz pulse generation by optical rectification in stoichiometric and cryo-cooled congruent lithium niobate. *J. Mod. Opt.* **62**, 1486–1493 (2015). doi: 10.1080/09500340.2013.868547
115. S.-W. Huang *et al*., High conversion efficiency, high energy terahertz pulses by optical rectification in cryogenically cooled lithium niobate. *Opt. Lett.* **38**, 796–798 (2013). doi: 10.1364/OL.38.000796; pmid: 23455302
116. J. A. L'huillier *et al*., Generation of THz radiation using bulk, periodically and aperiodically poled lithium niobate—Part 2: Experiments. *Appl. Phys. B* **86**, 197–208 (2007). doi: 10.1007/s00340-006-2528-z
117. H. T. Olgun *et al*., Highly efficient generation of narrowband terahertz radiation driven by a two-spectral-line laser in PPLN. *Opt. Lett.* **47**, 2374–2377 (2022). doi: 10.1364/OL.448457; pmid: 35561354
118. A. Rueda *et al*., Efficient microwave to optical photon conversion: An electro-optical realization. *Optica* **3**, 597 (2016). doi: 10.1364/OPTICA.3.000597
119. Y. Xu *et al*., Bidirectional interconversion of microwave and light with thin-film lithium niobate. *Nat. Commun.* **12**, 4453 (2021). doi: 10.1038/s41467-021-24809-y; pmid: 34294711
120. P. O. Weigel *et al*., Bonded thin film lithium niobate modulator on a silicon photonics platform exceeding 100 GHz 3-dB electrical modulation bandwidth. *Opt. Express* **26**, 23728–23739 (2018). doi: 10.1364/OE.26.023728; pmid: 30184869
121. F. Yang *et al*., Monolithic thin film lithium niobate electro-optic modulator with over 110 GHz bandwidth. *Chin. Opt. Lett.* **20**, 022502 (2022). doi: 10.3788/COL202220.022502
122. A. J. Mercante *et al*., Thin film lithium niobate electro-optic modulator with terahertz operating bandwidth. *Opt. Express* **26**, 14810–14816 (2018). doi: 10.1364/OE.26.014810; pmid: 29877417
123. Z. Y. Cheng, C. S. Tsai, Baseband integrated acousto-optic frequency shifter. *Appl. Phys. Lett.* **60**, 12–14 (1992). doi: 10.1063/1.107347
124. L. Shao *et al*., Integrated microwave acousto-optic frequency shifter on thin-film lithium niobate. *Opt. Express* **28**,

23728–23738 (2020). doi: 10.1364/OE.397138; pmid: 32752365
125. L. Shao *et al*., Microwave-to-optical conversion using lithium niobate thin-film acoustic resonators. *Optica* **6**, 1498 (2019). doi: 10.1364/OPTICA.6.001498
126. C. J. Sarabalis, T. P. McKenna, R. N. Patel, R. Van Laer, A. H. Safavi-Naeini, Acousto-optic modulation in lithium niobate on sapphire. *APL Photonics* **5**, 086104 (2020). doi: 10.1063/5.0012288
127. W. Jiang *et al*., Efficient bidirectional piezo-optomechanical transduction between microwave and optical frequency. *Nat. Commun.* **11**, 1166 (2020). doi: 10.1038/s41467-020-14863-3; pmid: 32127538
128. K. Buse, A. Adibi, D. Psaltis, Non-volatile holographic storage in doubly doped lithium niobate crystals. *Nature* **393**, 665–668 (1998). doi: 10.1038/31429
129. P. Ying *et al*., Low-loss edge-coupling thin-film lithium niobate modulator with an efficient phase shifter. *Opt. Lett.* **46**, 1478–1481 (2021). doi: 10.1364/OL.418996; pmid: 33720216
130. S. Yang *et al*., Low loss ridge-waveguide grating couplers in lithium niobate on insulator. *Opt. Mater. Express* **11**, 1366 (2021). doi: 10.1364/OME.418423
131. M. Li *et al*., Integrated pockels laser. *Nat Commun* **13**, 5344 (2022). doi: 10.1038/s41467-022-33101-6
132. V. Snigirev *et al*., Ultrafast tunable lasers using lithium niobate integrated photonics. arXiv:2112.02036 [physics. optics] (2021). doi: 10.48550/ARXIV.2112.02036
133. Y. Minet *et al*., Pockels-effect-based adiabatic frequency conversion in ultrahigh-Q microresonators. *Opt. Express* **28**, 2939–2947 (2020). doi: 10.1364/OE.378112; pmid: 32121971
134. M. Jankowski *et al*., Quasi-static optical parametric amplification. *OSA Nonlinear Optics* **2021**, NW3A.1 (2021). doi: 10.1364/NLO.2021.NW3A.1
135. C. Cochard, T. Spielmann, N. Bahlawane, A. Halpin, T. Granzow, Broadband characterization of congruent lithium niobate from mHz to optical frequencies. *J. Phys. D Appl. Phys.* **50**, 36LT01 (2017). doi: 10.1088/1361-6463/aa80b8
136. M. Lee, Dielectric constant and loss tangent in $LiNbO_3$ crystals from 90 to 147 GHz. *Appl. Phys. Lett.* **79**, 1342–1344 (2001). doi: 10.1063/1.1399305
137. A. Herter *et al*., Terahertz waveform synthesis from integrated lithium niobate circuits. arXiv:2204.11725 [physics. optics] (2022). doi: 10.48550/ARXIV.2204.11725
138. Q. Ma, G. Ren, K. Xu, J. Z. Ou, Tunable optical properties of 2D materials and their applications. *Adv. Opt. Mater.* **9**, 2001313 (2021). doi: 10.1002/adom.202001313
139. M. L. Bortz *et al*., Noncritical quasi-phase-matched second harmonic generation in an annealed proton-exchanged $LiNbO_3$ waveguide. *IEEE J. Quantum Electron.* **30**, 2953–2960 (1994). doi: 10.1109/3.362710
140. P. S. Kuo, Noncritical phasematching behavior in thin-film lithium niobate frequency converters. *Opt. Lett.* **47**, 54–57 (2022). doi: 10.1364/OL.444846; pmid: 34951881
141. D. Thomson *et al*., Roadmap on silicon photonics. *J. Opt.* **18**, 073003 (2016). doi: 10.1088/2040-8978/18/7/073003
142. M. Yu *et al*., Femtosecond pulse generation via an integrated electro-optic time lens. arXiv:2112.09204 [physics. optics] (2021). doi: 10.48550/ARXIV.2112.09204
143. T. Vanackere *et al*., "Micro-transfer printing of lithium niobate on silicon nitride" in *2020 European Conference on Optical Communications (ECOC)* (IEEE, 2020).
144. C. Wang *et al*., Ultrahigh-efficiency wavelength conversion in nanophotonic periodically poled lithium niobate waveguides. *Optica* **5**, 1438 (2018). doi: 10.1364/OPTICA.5.001438

ACKNOWLEDGMENTS

We thank G. Keeler, S. Diddams, and K. Vahala for discussions and assistance. We also thank many colleagues with whom we have had the privilege to work and learn about the ever-expanding field of LN photonics. We thank Y. Chen for assistance with graph preparations. We also thank J. Rutledge, T. Allison, X. Wu, Y. Li, B. Zhang, and A. Marandi for providing the data for the spectral graphs in Fig. 4. **Funding:** The material is based on work supported by the Australian Research Council (ARC) under awards ARC DP190101576, DP190102773, and DP220100488; the Defense Advanced Research Projects Agency (DARPA) under awards HR0011-20-2-0044, HR0011-15-C-0055, and DARPA-RA-18-02-YFA-ES-578; the National Science Foundation (NSF) under awards CCF-1918549, OIA-2040695, OMA-2137723, and OMA-2138174; the Department of Energy (DOE) under award DE-AC02-76SF00515; the Defense Threat Reduction Agency (DTRA) under award HDTRA11810047; the Beijing Natural Science Foundation under award Z220008; and the National Key Research and Development Program of China under award 2016QY01W0200. **Competing interests:** M.Z. and M.L. are involved in developing LN technologies at HyperLight Corporation and have an equity interest in HyperLight Corporation. M.Z. is an officer and member of the board of directors at HyperLight Corporation. M.Z. and M.L. are named on patents WO2018031916A1, US11092873B1, and WO2020172585A1. M.L. is named on patents US20210223657A1, US20210096444A1, and WO2022115392A1. C.L. and M.F. are named on patent US20220252958A1. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

Submitted 1 July 2022; accepted 3 November 2022
10.1126/science.abj4396

RESEARCH ARTICLES

OPTICS

Why optics needs thickness

David A. B. Miller*

This study shows why and when optical systems need thickness as well as width or area. Wave diffraction explains the fundamental need for area or diameter of a lens or aperture to achieve some resolution or number of pixels in microscopes and cameras. This work demonstrates that if we know what the optics is to do, even before design, we can also deduce the minimum required thickness. This limit comes from diffraction combined with a concept called overlapping nonlocality *C* that can be deduced rigorously from just the mathematical description of what the device is to do. *C* expresses how much the input regions for different output regions overlap. This limit applies broadly to optics, from cameras to metasurfaces, and to wave systems generally.

Modern micro- and nanofabrication methods allow for the creation of complex optics well beyond historic lenses, mirrors, and prisms, giving optics that does what we want, not just what previous optics offered. Such complex designs can, however, require long calculations and may be difficult to fabricate. The complexity also makes it hard to anticipate what may be possible. So, we want simple limits to guide us. What minimum sizes might we need, for example? From diffraction, we understand how the minimum width or area of the optics must grow in proportion with the resolvable spots or pixels. However, there has been no corresponding basic understanding of how thick the optics must be or even why optics fundamentally might require thickness.

In this work, I show why optics and other wave systems may need thickness and derive quantitative limits. Optics may need to be nonlocal—the output at some point may need to depend on the inputs at many positions. Such nonlocality means that we need to communicate sideways within the structure or system. If we only need one such communication channel, one thin layer may be enough. However, if the input position ranges for one output point need to overlap with those for another output point, we have overlapping nonlocality (ONL). A key realization is that this ONL leads to thickness in optics.

Here, I introduce ONL and define it as the required number *C* of such sideways communication channels. A basic result is that the ONL comes from just the mathematical specification of what the device is to do. We can calculate *C* quite rigorously before starting design. Then, with some heuristics from diffraction, we can deduce minimum thicknesses or cross-sectional areas for the optics from *C*. This approach gives limits for many optical components, including imagers and metasurface structures for many possible applications. More generally, it bounds sizes for complex wave systems of any kind, including radio-frequency and acoustic systems.

Two recent questions in metasurfaces motivated this work. First, can we shrink the distance between a lens and the output plane in an imager—i.e., "squeezing space" (*1*), possibly with a spaceplate (*2*–*5*)? Second, what kinds of mathematical operations could we perform—for example, on an image—using some metasurface structure with some thickness (*6*–*8*)? This approach gives meaningful answers to these questions and others. It gives limits even for operation at just one frequency—so it is complementary to a spaceplate bandwidth limit (*4*) from the amount of material in the device (*9*, *10*)—and to related semiempirical limits (*11*). It also complements other recent limits on maximal enhancement of material response (*12*) and minimum thicknesses for local functions, such as perfect absorption (*12*) or reflection (*13*). Limits in optics and electromagnetics are of increasing interest (*14*). The concepts and results in our approach may allow different directions in this field and may find applications in other areas with complex optics, such as mode converters (*15*–*17*) and optical networks (*18*) in neural (*19*–*21*) and other (*22*–*24*) processing and interconnects (*25*).

An optical system (Fig. 1A) routes the light from an input surface to an output surface. Adding a dividing surface that mathematically cuts through both the input and output surfaces defines a transverse aperture. A minimum area or thickness for this aperture can be deduced by counting the number *C* of independent channels that must pass through it. For a camera or imager, we can evaluate *C* intuitively. Singular value decomposition (SVD) gives a rigorous mathematical approach for optical and wave systems generally.

ONL for imaging systems

An imager might have a lens surface as its input and an array of pixel sensors as its output. We take it to be nominally lossless—other than for incidental losses, such as weak background absorption, minor surface roughness scattering, or reflection losses—as it routes essentially all the relevant input power to the outputs. We also presume reciprocal optics—if waves can flow in one direction, then their phase conjugates can flow in the reverse direction with the same transmission factor.

An imager takes a set of N overlapping orthogonal inputs and maps them, one by one, to its N separate output pixels (see supplementary text S1 for extended discussion and proofs). We presume, as is typical for an imager, that the input power for each output pixel is distributed essentially uniformly over the input surface.

We now divide both input and output surfaces mathematically in half with surface S in the y-z plane. Now, an imager is very much larger than a wavelength. So, we presume we can construct new approximate basis sets for each half of the input surface, assigning a number of basis functions in proportion to the area of each part—so, $N/2$ input basis functions for each half. We presume that, in combination, this new divided pair of basis sets is approximately still able to describe all the N orthogonal input functions.

Now consider the mapping between the right half of the input surface and the left half of the output surface (Fig. 1D). Although $N/2$ orthogonal basis functions are associated with the right half of the input surface, we expect that half of those will be associated with forming images on the right half of the output plane. So, only $C_{RL} = N/4$ channels are associated with transferring power from the right half of the input plane to pixels on the left half of the output plane. Similarly, a number $C_{LR} = N/4$ of left-to-right channels are needed for waves from the left input surface to the right output surface.

In deducing the total number C of channels that must pass from right to left through the transverse aperture, we might think that we could neglect any left-to-right channels because they are going in the other direction. However, by reciprocity, associated with those $C_{LR} = N/4$ left-to-right channels, there must also be an equal number of reciprocal or backward versions of those channels from the output pixels on the right to the input surface on the left. So, altogether, we must physically allow for

$$C = C_{RL} + C_{LR} \quad (1)$$

channels crossing the dividing surface from right to left (or from left to right). In what follows, Eq. 1 applies quite generally. (Theoretically, nonreciprocal optics could eliminate the backward channels, reducing C by up to a factor

Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA.
*Corresponding author. Email: dabm@stanford.edu

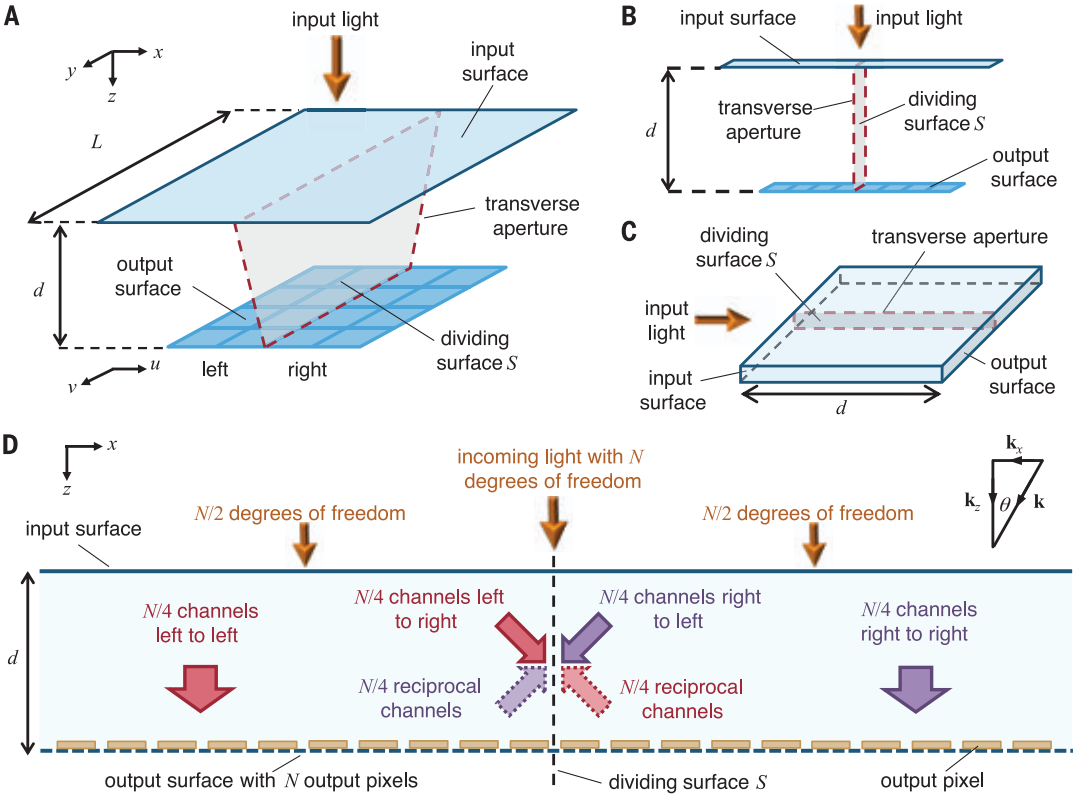


Fig. 1. Imaging systems and relevant surfaces and channels. (**A**) The input surface of an imaging system and the corresponding array of pixels on the output surface. We presume the surfaces are separated in z by some distance d. A dividing surface S that cuts through both input and output surfaces defines a transverse aperture. (**B** and **C**) A 1D imager, viewed either as a vertical slice that is thin in the y direction and has thickness d (B) or as a thin slab in the y direction with length d in x (C), as in a photonic integrated circuit. (**D**) Required internal channels when dividing an imaging system with a large number N of pixels and degrees of freedom into two equal parts.

of 2; see supplementary text S1). So, for our imager

$$C = N/4 + N/4 = N/2 \quad (2)$$

Note, C here comes from how an imager must work and the number of pixels, not from any specific design or size of the imager.

At this point, we can formally define ONL and C. The ONL C associated with a dividing surface S passing through the input and output surfaces is the number C of orthogonal channels that must cross from inputs on one side of S to outputs on the other side of S to implement the desired optical function, summing over both directions (left to right and right to left) of flow.

We emphasize that nonlocality itself does not require multiple channels. A single-mode optical fiber can have multiple taps spaced as far apart as we like. Appropriate light into those taps could be coherently combined to emerge at the fiber end, giving a very nonlocal system with only one channel. Rather, it is the overlapping nature of nonlocality in some systems—where different output points require different combinations of the light at some of the same input points—that necessitates multiple channels.

Required area or thickness of the transverse aperture

We presume that the optical systems of interest are sufficiently nonlocal that they require propagation of these C channels over transverse distances of many wavelengths in connecting input and output points (so we exclude any nearly local system, such as a very local differentiation just comparing inputs <1 wavelength apart). So, we presume propagating electromagnetic waves for these channels, not evanescent fields or near-field electromagnetic terms (*16*). We presume simple local dielectrics—the polarization at some point depends just on the field at that point—so we neglect any nonlocality from plasmons or other compound excitations. Thus, we can use wave diffraction heuristics to predict size limits. For simplicity, we effectively consider just one electromagnetic polarization, but the same results would apply to each polarization.

We start by pretending that the space between the input and output surfaces contains a uniform dielectric of refractive index n_r with light of free-space wavelength λ_o. Diffraction heuristics (supplementary text S2) tell us that in a narrow slit aperture, as in Fig. 1B, the maximum number of channels through the aperture corresponds to one for every $\lambda_o/2n_r$ of distance in the z direction. If this space is a nonuniform dielectric with maximum refractive index n_{max}, we conjecture at least $\lambda_o/2n_{max}$ per channel. Note, such a conjecture does not prove that no structure could do better. It is, however, consistent with typical behavior in the number of modes supported by slab waveguides, for example.

Practically, we may be limited to using only some fraction of the full 180° range of angles inside the structure—equivalently, a fraction α (≤1) of the available k-space (i.e., of the component k_z, as in Fig. 1D)—reducing the available channels proportionately. For example, if the internal angle is restricted to a range 0 to θ (Fig. 1D), then $\alpha = 1 - \cos\theta$. Hence, we conjecture in this one-dimensional (1D) case that we need a thickness

$$d \geq \frac{C\lambda_o}{2\alpha n_{max}} \quad (3)$$

We can extend this heuristic argument to the area A of a 2D transverse aperture, as in Fig. 1A, proposing

$$A \geq C\frac{1}{\alpha^2}\left(\frac{\lambda_o}{2n_{max}}\right)^2 \quad (4)$$

where we regard α^2 as the fraction of the 2D k_x,k_z k-space that we are practically able to use in design. Equation 4 is equivalent to an area of at least $(\lambda_o/2\alpha n_{max})^2$ for each channel through the transverse aperture.

Minimum thicknesses for imagers and related optical systems

We now apply Eqs. 3 and 4 to imagers. For a 1D imager with N_x pixels in a horizontal line in the x direction, as in Fig. 1, B and C, from Eq. 2, we have $C = N_x/2$, so from Eq. 3

$$d \geq \frac{N_x\lambda}{4\alpha n_{max}} \quad (5)$$

For a 2D imager, as in Fig. 1A, with N pixels (so $C = N/2$ from Eq. 2) and some characteristic width or diameter L, with transverse aperture area $A \sim Ld$, then from Eq. 4

$$d \geq \frac{N}{2L}\frac{1}{\alpha^2}\left(\frac{\lambda}{2n_{\max}}\right)^2 \quad (6)$$

To exploit the transverse aperture area in Eqs. 4 or 6 effectively, we may need to interleave degrees of freedom originally in x into the y dimension in the transverse aperture. This dimensional interleaving (DI) (supplementary text S3) is possible in optics, and we can design supercouplers to achieve it (supplementary text S4), including devising limits for these. Many approaches to optics, including free-space propagation, conventional imaging systems, simple dielectric stack structures, and 2D photonic crystals, do not, however, appear to support DI. In such cases, the thickness of these 2D systems may end up as the 1D limit (Eqs. 3 and 5).

We compare these limits on d with specific designs for imagers and spaceplates in supplementary text S5, showing that these limits are both obeyed and approached in existing optimized designs. A 12-megapixel smartphone camera (*26*) would require >~1.7-mm thickness if designed with typical (*26*) (<45°) maximum ray angles, even with no thickness in the lenses themselves (i.e., within about a factor of 3 of actual ~5-mm smartphone camera thicknesses). The multilayer spaceplate design in (*5*) has a designed thickness of 44.6 wavelengths, which is quite close to the predicted limit of 30 wavelengths.

An imager is a space-variant system—it looks different at different positions in the input or output. Several other such systems, such as Fourier transformers (*27*), mode sorters (*17*), and connection networks (*18*), can be analyzed similarly (supplementary text S6).

ONL for general linear optical devices

An imager or mode sorter has a pixelated output, which simplifies counting. Many optical devices, however, have no such pixelation, with continuous functions on input and output surfaces. The kernel—the linear operator relating the field at output points to that at input points—may be more local than the imager's global kernel; a spatial differentiator, which also may not be unitary, relates an output region to a small number of adjacent input regions (Fig. 2). The kernel may not be symmetric left to right, and it may not be obvious where to put the dividing surface. Fortunately, a SVD (*16*) approach is both compatible with the arguments so far and with these other cases.

With coordinates x and y on the input face and u and v on the output face (Fig. 1), as in the formalism of (*8*), generally

$$\Phi(u,v) = \iint D(u,v;x,y)\Psi(x,y)dxdy \quad (7)$$

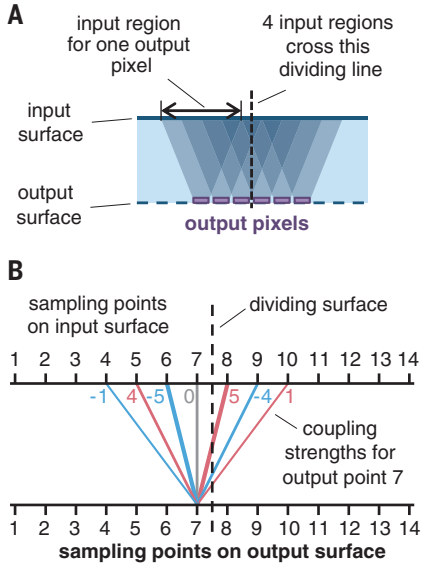

A
input region for one output pixel
4 input regions cross this dividing line
input surface
output surface
output pixels
B
sampling points on input surface
dividing surface
1 2 3 4 5 6 7 8 9 10 11 12 13 14
-1 4 -5 0 5 -4 1
coupling strengths for output point 7
1 2 3 4 5 6 7 8 9 10 11 12 13 14
sampling points on output surface

Fig. 2. Connections between input and output pixels. (**A**) A general example with an ONL of $C = 4$. (The trapezoids show which pixel is connected to which of the overlapping input regions.) (**B**) The coupling strengths between the input sampling points and the output sampling point 7 for a central finite-difference approximation to a fifth-order derivative. Coupling strengths for the other output points are shifted sideways as appropriate, as in (A).

where $D(u,v;x,y)$ is the kernel or the device operator (*15*, *16*), relating the output function $\Phi(u,v)$ to the input function $\Psi(x,y)$.

Choosing a dividing surface at input and output positions x_o and u_o, respectively, we have a divided operator D_{RL} restricted to the right part of the input and the corresponding left part of the output

$$D_{RL}(u,v;x,y) = \Theta(x-x_o)\Theta(u_o-u)D(u,v;x,y) \equiv \begin{cases} D(u,v;x,y) & u \leq u_o,\ x \geq x_o \\ 0 & u > u_o \\ 0 & x < x_o \end{cases} \quad (8)$$

where $\Theta(z)$ is the Heaviside (or step) function.

To find C, we start by finding the SVD of $D_{RL}(u,v;x,y)$. [Technically, we are establishing the necessary mode converter basis sets (*16*) to implement this right-to-left operator.] We then decide how many of the singular values (i.e., coupling strengths) have a magnitude above some small threshold and use that as the number of required right-to-left channels, C_{RL}. If necessary, from a corresponding left-to-right operator

$$D_{LR}(u,v;x,y) = \Theta(x_o-x)\Theta(u-u_o)D(u,v;x,y) \quad (9)$$

we similarly deduce a practical number of left-to-right channels C_{LR}. As above from Eq. 1, we add C_{RL} and C_{LR} to obtain C. For symmetric kernels, we may only need to calculate either C_{RL} or C_{LR} and double it.

If, for some kernel, it is not obvious where to put the dividing surface, we could repeat the calculation for all reasonable choices of dividing surface positions and choose the largest result for C. We should, however, keep the output central point u_o beneath its corresponding input range (supplementary text S7).

Constructing matrices for general linear optical devices

Because any such device operator D in a real physical system gives finite output for finite input, it is necessarily a Hilbert-Schmidt operator and hence is also compact (*16*, *28*). This means that it can be represented to any precision by a sufficiently large matrix $\mathbf{D}$.

The matrix elements of $\mathbf{D}$ are the couplings between specific chosen sampling points for the functions in the input and output spaces. Matrices $\mathbf{D}_{RL}$ and $\mathbf{D}_{LR}$ are then just truncated versions of $\mathbf{D}$; for example, for a 1D problem, $\mathbf{D}_{RL}$ and $\mathbf{D}_{LR}$ are just the upper-right and lower-left quadrants of $\mathbf{D}$. Standard matrix algebra gives the SVD of $\mathbf{D}_{RL}$ and, if necessary, that of $\mathbf{D}_{LR}$, which allows us to deduce C from Eq. 1. For pixelated optics, we could choose sampling points in the middle of each such pixel; essentially, we are then deducing limiting sizes for the optics so that it could give the right fields at least at these points.

For continuous functions and/or those without pixelation, we could just choose points with a separation that is close enough—intuitively sufficient to represent even the smallest bump in the function. The criterion for "close enough" is then that the number of singular values of the resulting matrices, above some chosen small threshold of relative magnitude, has converged. Increasing the density of sampling points beyond this makes essentially no difference to the resulting C. Generally, experience in the SVD description of optics (*16*) shows this behavior quite consistently, with convergence guaranteed by the operator compactness and associated sum rules (*16*, *28*).

In space-invariant optics (*8*), where the behavior depends only on the relative separation of input and output points, we are then convolving with a fixed kernel. Then, D simplifies (*8*) to

$$D(u,v;x,y) \rightarrow D(x-v,y-v) \quad (10)$$

Because the absolute position no longer matters, we simply choose one specific position for the calculations—e.g., for the output, such as $u = 0$, $v = 0$—and evaluate the matrices as required.

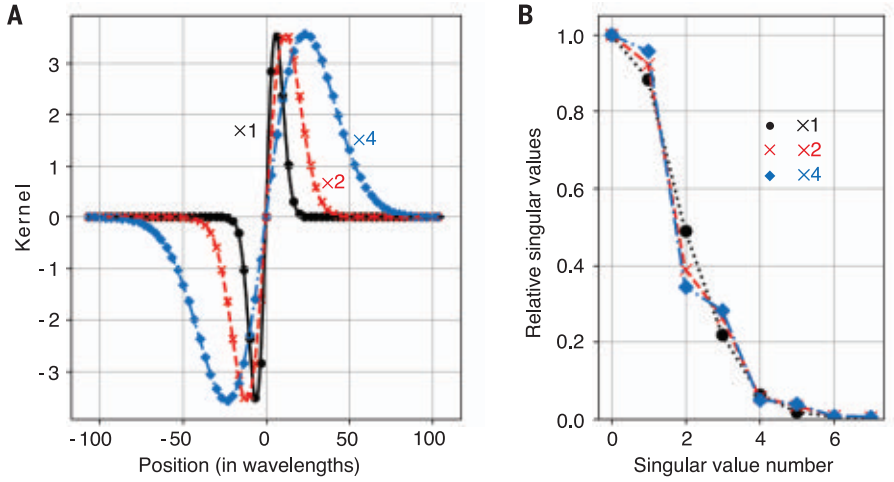


Fig. 3. An "*x* times Gaussian" kernel at three scales. (**A**) The kernels ×1—the original scale—(β = 1) (circles and solid line), ×2 larger (β = 2) (crosses and dashed line), and ×4 larger (β = 4) (diamonds and dot-dashed line). The points correspond to effective sampling points for a numerical aperture NA = 0.15. (**B**) The corresponding relative magnitude of the singular values (including from left-to-right and right-to-left matrices) for the three scales of kernels (with lines as a guide), using symbols and colors as in (A).

Much such metasurface discussion uses k-space (or Fourier) representations of functions and (space-invariant) kernels (*1*, *5*, *6*); pixels are not explicitly used. k-values must be smaller than $k = 2\pi n_r/\lambda_o$ for a propagating wave in the background material with refractive index n_r, or a smaller maximum value $k_{x\max} = 2\pi\mathrm{NA}/\lambda_o$ if the input and output optics has a finite numerical aperture NA. In this case, we can use a sampling theory approach to get effective spatial sampling points (supplementary text S8). With N sampling points in one dimension, these are spaced by

$$\delta l = \frac{\lambda_o}{2\mathrm{NA}} \equiv \frac{L}{N} \tag{11}$$

where $L = N\delta l$ now becomes the nominal width of the surfaces for this calculation.

Example calculations of ONL

Pixelated systems

Consider a device implementing a centered finite-difference fifth-order linear derivative (*29*) in the x direction, in the spirit of Fig. 1B. Seven adjacent, equally spaced sampling points would have weights proportional to –1, 4, –5, 0, 5, –4, and 1, as sketched in Fig. 2B for a dividing surface between points 7 and 8.

The connections between input points on the right of the dividing surface and output points on the left are expressed by the 3-by-3 matrix

$$\mathbf{D}_{\mathrm{RL}} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 5 & -4 & 1 \end{bmatrix} \tag{12}$$

This matrix contains the connections from input points 8, 9, and 10 (corresponding to matrix columns) on the right to output points 5, 6, and 7 (corresponding to matrix rows) on the left. All other connections across the dividing surface are zero. (See supplementary text S9 and fig. S7 for the full matrix $\mathbf{D}$ and matrices $\mathbf{D}_{\mathrm{RL}}$ and $\mathbf{D}_{\mathrm{LR}}$.)

A standard numerical linear algebra calculation of the SVD of $\mathbf{D}_{\mathrm{RL}}$ (using the numpy Python library) gives the three singular values 7.568, 1.684, and 0.080, so $C_{\mathrm{RL}} = 3$. If we similarly analyze the connections from left to right across the dividing surface, from input points 5, 6, and 7 to output points 8, 9, and 10 with this antisymmetric kernel, the resulting matrix ends up being $\mathbf{D}_{\mathrm{LR}} = -\mathbf{D}_{\mathrm{RL}}^T$ (fig. S7) and has the same set of three singular values, giving $C_{\mathrm{LR}} = 3$. So, for this fifth-order finite difference derivative, we require $C = C_{\mathrm{RL}} + C_{\mathrm{LR}} = 6$ (as in Eq. 1).

In this SVD, we see an important behavior that we exploit later: Not all the required channels are equally strong, and some may be negligible or nearly so. In fact, the third channel in each direction is nearly 100 times as weak as the first (0.080 compared with 7.568), which suggests that, if we only need a moderately good approximation for our derivative, we might need only two channels in each direction (so $C = 4$).

We can apply the same approach for other pixelated systems; finite impulse response filters and discrete wavelets, such as Daubechies wavelets, give additional examples (supplementary text S10). For pixelated systems, in some simple cases, it is quite straightforward to understand ONL intuitively in optics (supplementary text S11).

Continuous systems

As an example with continuous functions and kernel, we use

$$D(u;x) = \frac{(x-u)}{\beta}\exp\left[-\frac{(x-u)^2}{\beta^2\Delta_t^2}\right] \tag{13}$$

which is a real, 1D version of the "x times Gaussian" ∂_x kernel from (*7*) that gives a smoothed differentiation. We allow for a scale-up factor β by which we can increase the distance scale of the kernel, with β = 1 corresponding directly to (*7*). As in (*7*), we take $\Delta_t \simeq 8.325$ wavelengths and NA = 0.15, which, by Eq. 11, leads to sampling points spaced by ~*3.33* wavelengths. The resulting kernels for three different scales are shown in Fig. 3 together with the corresponding sets of relative singular values, including both right-to-left and left-to-right singular values in the same graph.

The total number of singular values equals the number of sampling points. After the first several singular values, however, the magnitudes of the remaining ones fall off very rapidly (we plot only the first eight in this work). Also, the set of strongly coupled singular values is essentially the same for all three scales of kernel. Once we have a large number of sampling points over the range where the kernel function is changing substantially, the relative size of the singular values converges. This illustrates that the ONL C is a property of the form of the function, not its scale, at least beyond some practical minimum scale. In all three cases shown, only the first six singular values have a relative size >0.01. So, practically, we might choose $C = 6$ for this function.

Thicknesses for space-invariant kernels

These examples show many interesting, discrete, and continuous space-invariant kernels and operations that could be performed with values of C from ~4 to ~8. Such numbers are likely still large enough that Eqs. 3 and 4 are usable at least as a first guide. (More sophisticated approaches using SVD are possible for thin structures and/or small C without relying on the heuristics behind Eqs. 3 and 4; see supplementary text S12.) Even without DI, such kernels might be implemented practically in structures that, for optical and near-infrared wavelengths, are only some small number of micrometers thick. A comparison with the "x times Gaussian" kernel design in (*7*) shows that, with an ~6-wavelength thickness, it also exceeds the minimum required thickness of ~2 wavelengths (supplementary text S13).

Discussion

These examples over a wide range of situations with waves, including pixelated, continuous, space-variant, and space-invariant systems, show that we have a general method to predict

fundamental minimum required thicknesses. The complete process is summarized and some additional discussion is provided in supplementary text S14 and S15, respectively. As illustrated above, systems with optimized designs already approach these limits within some small factor (e.g., a factor of 3 or less).

Flat optical systems offer many interesting possibilities for reducing the thickness of systems—e.g., by using metasurfaces to eliminate most of the thickness of lenses or other elements. This work shows that, especially for applications with large ONL, although we still need thickness overall, much of this thickness just needs to transport optical channels laterally; it may only need to be empty, uniform, or relatively simple waveguiding space, which would simplify overall system design.

REFERENCES AND NOTES

1. C. Guo, H. Wang, S. Fan, *Optica* **7**, 1133–1138 (2020).
2. O. Reshef *et al*., *Nat. Commun.* **12**, 3512 (2021).
3. J. T. R. Pagé, O. Reshef, R. W. Boyd, J. S. Lundeen, *Opt. Express* **30**, 2197–2205 (2022).
4. K. Shastri, O. Reshef, R. W. Boyd, J. S. Lundeen, F. Monticone, *Optica* **9**, 738–745 (2022).
5. A. Chen, F. Monticone, *ACS Photonics* **8**, 1439–1447 (2021).
6. A. Silva *et al*., *Science* **343**, 160–163 (2014).
7. H. Wang *et al*., *ACS Photonics* **9**, 1358–1365 (2022).
8. A. Overvig, A. Alù, *Laser Photonics Rev.* **16**, 2100633 (2022).
9. D. A. B. Miller, *J. Opt. Soc. Am. B* **24**, A1–A18 (2007).
10. D. A. B. Miller, *Phys. Rev. Lett.* **99**, 203903 (2007).
11. M. Gerken, D. A. B. Miller, *Appl. Opt.* **44**, 3349–3357 (2005).
12. Z. Kuang, L. Zhang, O. D. Miller, *Optica* **7**, 1746–1757 (2020).
13. M. I. Abdelrahman, F. Monticone, How Thin and Efficient Can a Metasurface Reflector Be? Universal Bounds on Reflection for Any Direction and Polarization. arXiv:2208.05533 [physics.optics] (2022).
14. P. Chao, B. Strekha, R. Kuate Defo, S. Molesky, A. W. Rodriguez, *Nat. Rev. Phys.* **4**, 543–559 (2022).
15. D. A. B. Miller, *Opt. Express* **20**, 23985–23993 (2012).
16. D. A. B. Miller, *Adv. Opt. Photonics* **11**, 679–825 (2019).
17. N. K. Fontaine *et al*., *Nat. Commun.* **10**, 1865 (2019).
18. S. Pai *et al*., *IEEE J. Sel. Top. Quantum Electron.* **26**, 1–13 (2020).
19. B. J. Shastri *et al*., *Nat. Photonics* **15**, 102–114 (2021).
20. F. Ashtiani, A. J. Geers, F. Aflatouni, *Nature* **606**, 501–506 (2022).
21. X. Lin *et al*., *Science* **361**, 1004–1008 (2018).
22. W. Bogaerts *et al*., *Nature* **586**, 207–216 (2020).
23. G. Wetzstein *et al*., *Nature* **588**, 39–47 (2020).
24. D. A. B. Miller, *Photon. Res.* **1**, 1–15 (2013).
25. D. A. B. Miller, *J. Lightwave Technol.* **35**, 346–396 (2017).
26. V. Blahnik, O. Schindelbeck, *Adv. Opt. Technol.* **10**, 145–232 (2021).
27. J. W. Goodman, *Introduction to Fourier Optics* (Macmillan, 2017).
28. D. A. B. Miller, An introduction to functional analysis for science and engineering. arXiv:1904.02539 [math.FA] (2019).
29. B. Fornberg, *Math. Comput.* **51**, 699–706 (1988).

ACKNOWLEDGMENTS

The author acknowledges stimulating conversations with S. Fan. **Funding:** D.A.B.M. received funding for this work from the MURI program supported by AFOSR grant no. FA9550-21-1-0312. **Competing interests:** The author declares that he has no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade3395
Supplementary Text
Figs. S1 to S11
References (*30*–*40*)

Submitted 9 August 2022; accepted 31 October 2022
10.1126/science.ade3395

IMMUNOLOGY

Past history of obesity triggers persistent epigenetic changes in innate immunity and exacerbates neuroinflammation

Masayuki Hata[1,2], Elisabeth M. M. A. Andriessen[3], Maki Hata[1], Roberto Diaz-Marin[2], Frédérik Fournier[2], Sergio Crespo-Garcia[1,2], Guillaume Blot[1,2], Rachel Juneau[1], Frédérique Pilon[1], Agnieszka Dejda[1], Vera Guber[1], Emilie Heckel[4], Caroline Daneault[5], Virginie Calderon[6], Christine Des Rosiers[5], Heather J. Melichar[7], Thomas Langmann[8], Jean-Sebastien Joyal[4], Ariel M. Wilson[1], Przemyslaw Sapieha[1,2]*

Age-related macular degeneration is a prevalent neuroinflammatory condition and a major cause of blindness driven by genetic and environmental factors such as obesity. In diseases of aging, modifiable factors can be compounded over the life span. We report that diet-induced obesity earlier in life triggers persistent reprogramming of the innate immune system, lasting long after normalization of metabolic abnormalities. Stearic acid, acting through Toll-like receptor 4 (TLR4), is sufficient to remodel chromatin landscapes and selectively enhance accessibility at binding sites for activator protein-1 (AP-1). Myeloid cells show less oxidative phosphorylation and shift to glycolysis, ultimately leading to proinflammatory cytokine transcription, aggravation of pathological retinal angiogenesis, and neuronal degeneration associated with loss of visual function. Thus, a past history of obesity reprograms mononuclear phagocytes and predisposes to neuroinflammation.

Age-related macular degeneration (AMD) is a multifactorial neuroinflammatory disease of the aging eye that is caused by genetic and environmental risk factors (*1*, *2*). It is the leading cause of irreversible blindness in the world (*3*). Investigation of the genetic predisposition to AMD identified variants in susceptibility gene loci but not causative gene mutations. Therefore, environmental factors likely contribute to the pathogenesis of AMD. This notion is supported by monozygotic twin studies demonstrating that, along with age, environmental factors determine disease susceptibility (*4*).

In the early stages of AMD, insoluble extracellular deposits called drusen, containing lipids, proteins, hydroxyapatite, and trace metals, form in the subretinal pigment epithelium (RPE) space (*5*–*7*). Components of drusen also include various inflammatory factors such as immunogenic modified lipoproteins, complement proteins, vitronectin, amyloid proteins, and immunoglobulins (*8*), which attract and activate cells of the innate immune system. These cells include resident retinal microglia and recruited monocytes or macrophages (*9*, *10*), which attempt to achieve tissue repair through low-grade inflammation. However, similarly to other neurodegenerative diseases, early AMD is triggered by chronic parainflammation, overly activated phagocytes, and dysregulation of the complement system (*11*, *12*). Moreover, inadequate clearance of drusen by phagocytes precipitates further pathologic inflammatory and angiogenic responses that lead to tissue damage that predisposes to late AMD (*13*). Late AMD is classified into two forms, neovascular AMD (nAMD) and non-neovascular AMD (geographic atrophy). nAMD causes >80% of vision loss in patients with AMD (*14*) and is characterized by the growth of abnormal blood vessels from the choroid under the macula, known as choroidal neovascularization (CNV). CNV progression leads to retinal edema or hemorrhage, culminating in photoreceptor death. Abnormal neovascularization ultimately causes fibrovascular scarring and can lead to permanent loss of central vision (*15*). In addition to a local immune response, progression of AMD is also influenced by heightened systemic inflammation (*16*). How distal inflammation influences AMD remains incompletely understood.

Among modifiable factors, obesity is the second most important risk factor after smoking for late AMD (*17*). Patients with nAMD have significantly larger volumes of visceral adipose tissue (*18*), and each increase of 0.1 in

[1]Department of Ophthalmology, Maisonneuve-Rosemont Hospital Research Centre, University of Montreal, Montreal, Quebec H1T 2M4, Canada. [2]Department of Biochemistry and Molecular Medicine, Maisonneuve-Rosemont Hospital Research Centre, University of Montreal, Montreal, Quebec H1T 2M4, Canada. [3]Department of Biomedical Sciences, Maisonneuve-Rosemont Hospital Research Centre, University of Montreal, Montreal, Quebec H1T 2M4, Canada. [4]Departments of Pediatrics, Ophthalmology, and Pharmacology, Centre Hospitalier Universitaire Ste-Justine Research Center, Montreal, Quebec H3T 1C5, Canada. [5]Department of Nutrition, University of Montreal, Montreal, Quebec, Plateforme métabolomique de l'Institut de Cardiologie de Montréal, Montreal, Quebec H3C 3J7, Canada. [6]Bioinformatics & Molecular Biology Core Facility, Institut de Recherches Cliniques de Montréal, Montreal, Quebec H2W 1R7, Canada. [7]Department of Medicine, Maisonneuve-Rosemont Hospital Research Centre, University of Montreal, Montreal, Quebec H1T 2M4, Canada. [8]Laboratory for Experimental Immunology of the Eye, Department of Ophthalmology, Faculty of Medicine and University Hospital Cologne, University of Cologne, 50931 Cologne, Germany.
***Corresponding author. Email: mike.sapieha@umontreal.ca**

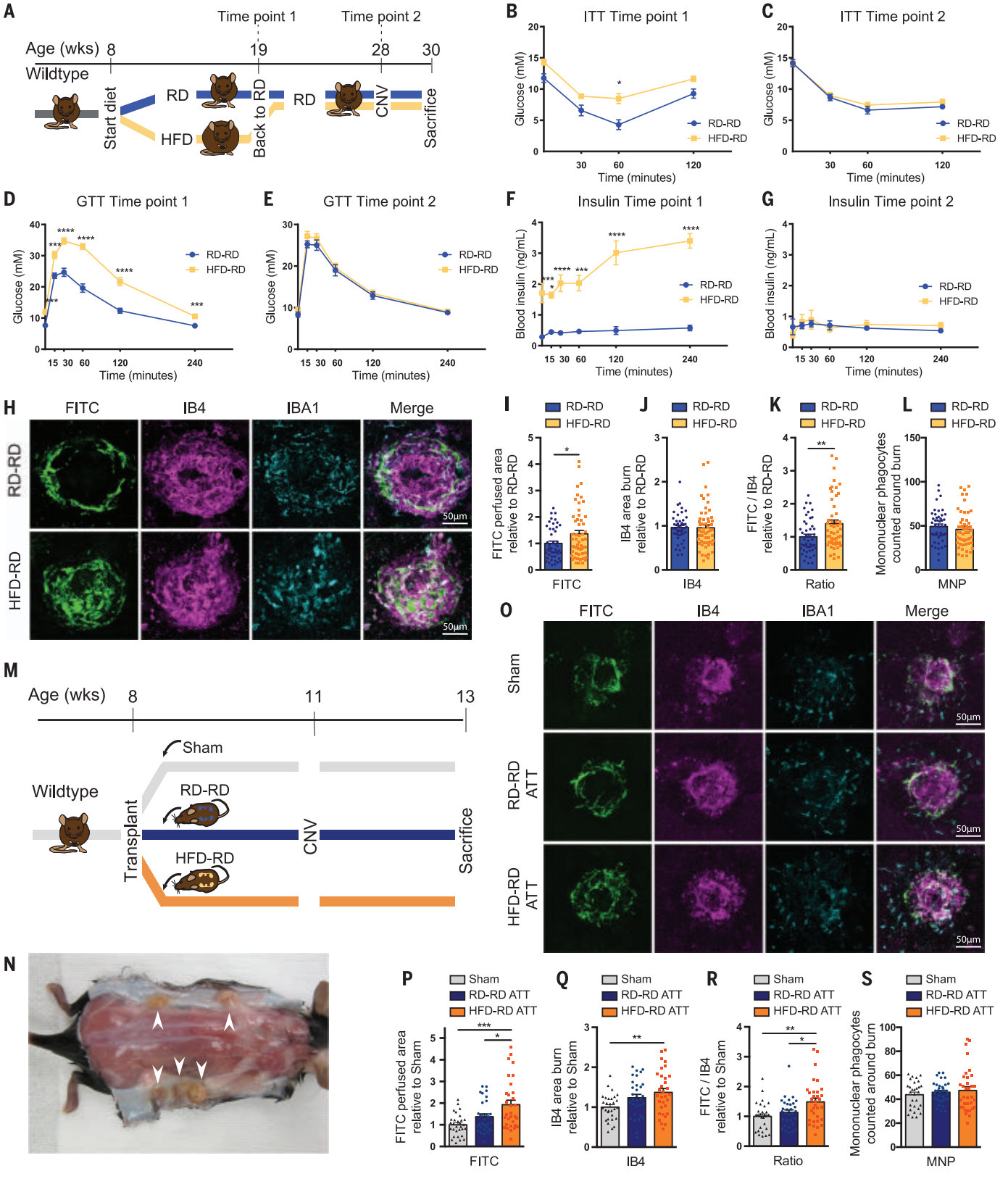


Fig. 1. DIO triggers long-term changes in eWAT that exacerbate pathological angiogenesis. (**A**) Experimental schematic in which mice started a HFD at 8 weeks of age and were then switched back to a RD at 19 weeks, which is time point 1; this group is HFD-RD. Control mice were fed a RD throughout and are referred to as RD-RD. At 28 weeks, which is time point 2, mice were subjected to laser-induced CNV. Two weeks after CNV induction, mice were euthanized and eyes collected. (**B** and **C**) ITT on RD-RD and HFD-RD mice at time point 1 (RD-RD, n = 5; HFD-RD, n = 6) (B) and time point 2 (RD-RD, n = 10; HFD-RD, n = 11) (C). (**D** and **E**) GTT on RD-RD and HFD-RD mice at time point 1 (RD-RD, n = 20; HFD-RD, n = 21) (D) and time point 2 (RD-RD, n = 18;

HFD-RD, *n* = 21) (E). (**F** and **G**) Plasma insulin levels in RD-RD and HFD-RD mice at time point 1 (RD-RD, *n* = 4; HFD-RD, *n* = 5) (F) and time point 2 (RD-RD, *n* = 6; HFD-RD, *n* = 5) (G) during GTT. (**H**) Representative confocal images of IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in RD-RD and HFD-RD mice. Scale bar, 50 μm. (**I** to **K**) Quantification of the FITC-dextran–labeled CNV area (I), the IB4-stained laser impact area (J), and the ratio of FITC/IB4 per laser burn (K) relative to RD-RD mice at D14; *n* = 46 burns for RD-RD; *n* = 54 burns for HFD-RD. (**L**) Number of IBA1$^+$ MNPs around the laser impact area at D14; *n* = 46 burns for RD-RD; *n* = 56 burns for HFD-RD. (**M**) Experimental schematic of the ATT in which recipient mice were transplanted with eWAT fat pads at 8 weeks of age from either RD-RD or HFD-RD donor mice. Sham surgeries of controls were performed similarly but without ATT. Mice were subjected to laser burns at 11 weeks and euthanized at week 13. (**N**) Adipose tissue grafts 3 weeks after ATT. The grafts (white arrowheads) were dissected from the epididymis of the donor mice and inserted in the backs of the recipient mice. Note the blood vessels that sprouted and are now supplying the grafts. (**O**) Compilation of representative compressed Z-stack confocal images showing IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in sham-operated mice and RD-RD ATT and HFD-RD ATT recipient mice. Scale bar, 50 μm. (**P** to **R**) Quantification of the FITC-dextran–labeled CNV area (P), the IB4-stained laser impact area (Q), and the ratio of FITC/IB4 per laser burn (R) relative to sham at day 14; *n* = 30 burns for sham, *n* = 31 burns for RD-RD ATT, and *n* = 32 burns for HFD-RD ATT. (**S**) Number of IBA1$^+$ MNPs around the laser impact area at day 14 ; *n* = 29 burns for sham, *n* = 35 burns for RD-RD ATT, and *n* = 34 burns for HFD-RD ATT. Data information: Comparisons between groups were analyzed using two-way ANOVA with Sidak's multiple-comparisons test [(B) to (G)], Student's unpaired *t* test [(I) to (L)], or one-way ANOVA with Tukey's multiple-comparisons test [(P) to (S)]. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Error bars represent mean ± SEM.

the waist/hip ratio, a measure of abdominal obesity, is associated with a 75% increase in the odds of late AMD (*17*). Longitudinal analysis from the Age-Related Eye Disease Study (AREDS) showed that a higher body-mass index (BMI > 30) was associated with progression to late AMD and more years of disease (*19*). The mechanisms through which obesity predisposes to late AMD remain poorly defined.

Compounded effects of obesity throughout life have been heavily investigated, revealing that weight loss in obese patients reduced adipose tissue inflammation and reinstated glycemic control (*20*–*23*). However, the long-term impact of prior obesity on the immune response in later life remains unknown. We sought to determine the long-term consequences of a past history of obesity on neuroinflammatory complications of the retina that lead to CNV and retinal degeneration and whether weight loss restores immune homeostasis in the aging eye.

RESULTS

Past history of obesity triggers persistent changes in visceral adipose tissue and predisposes mice to pathological angiogenesis in the retina

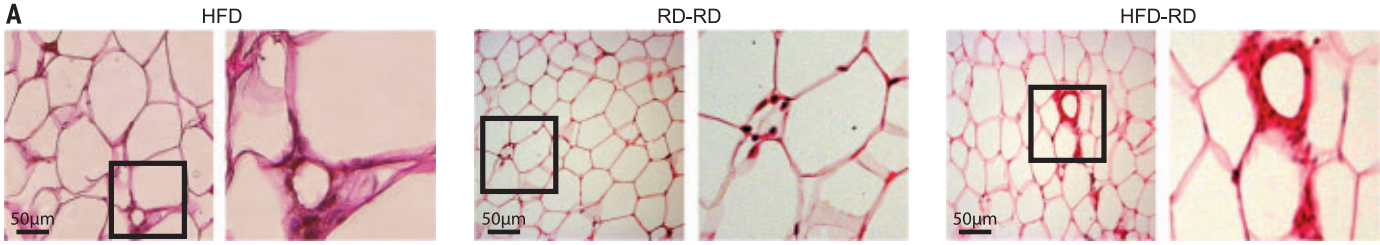
Diet-induced obesity (DIO) promotes systemic inflammation (*24*). Persistent obesity exacerbates CNV through activation of systemic innate immunity (*25*, *26*). However, whether weight loss after obesity can reverse this effect is poorly understood. We therefore set up weight gain–weight loss (WGWL) experiments in which we placed male C57BL/6J mice on a high-fat diet (HFD: 60% lipid content) for 11 weeks to provoke DIO and then switched them to a regular diet (RD: 10% lipid content, with lipid source as in the HFD) for 9 weeks to induce weight loss (HFD-RD mice). Control groups were fed a RD throughout the study (20 weeks; RD-RD mice) (Fig. 1A). After 11 weeks, mice on HFD gained three times more weight than mice on RD (fig. S1A, time point 1). The weight of HFD-RD mice gradually decreased after return to RD, returning to weights similar to RD-RD mice within 6 weeks and remaining similar throughout the course of study (fig. S1A, time point 2). The systemic metabolic consequences of DIO in HFD mice were normalized, with subsequent weight loss upon return to RD, as demonstrated by a comparable response to insulin challenge [insulin tolerance test (ITT); Fig. 1, B and C, and fig. S1, B and C] and glucose challenge [glucose tolerance test (GTT); Fig. 1, D and E]. Plasma insulin concentrations during GTT (Fig. 1, F and G) at time point 2 were similar between HFD-RD and RD-RD control mice.

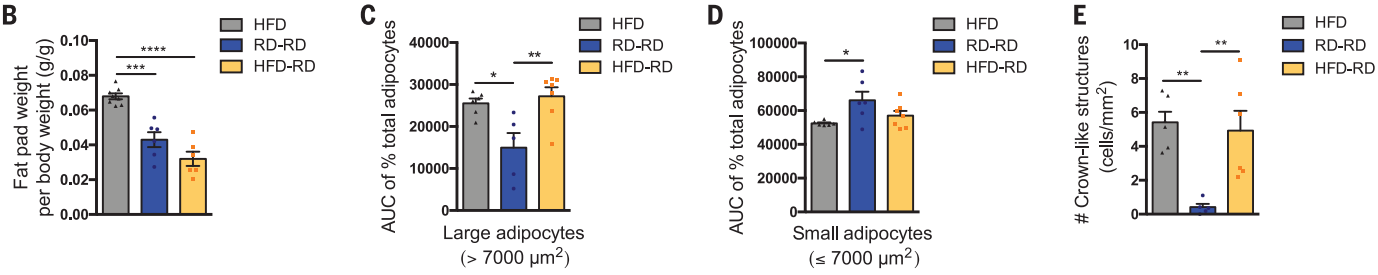
After 20 weeks, RD-RD mice and HFD-RD mice were subjected to laser-induced photocoagulation in the eye to trigger CNV (*27*) (Fig. 1A). Two weeks after treatment, CNV lesions and laser-burned areas were quantified with high molecular weight fluorescein isothiocyanate (FITC)–dextran and isolectin B4 (IB4) staining in choroid flatmounts. FITC-dextran perfusion permits visualization of neovessels with lumen, and IB4 stains endothelial cells in neovascularization and choriocapillaris beneath the laser-burned Bruch's membrane. Quantification of FITC-dextran–perfused newly formed vessels revealed a 40% increase in CNV in HFD-RD mice compared with that of RD-RD mice (Fig. 1, H to K). The average size of IB4-labeled lesion areas did not differ between groups, suggesting that the observed effect is directly on the nascent vasculature (Fig. 1, J and K). The numbers of mononuclear phagocytes [(MNPs), which are labeled with ionized calcium-binding adaptor molecule 1 (IBA1)] recruited to sites of injury in RPE-choroid-sclera complexes was comparable between RD-RD mice and HFD-RD mice (Fig. 1, H and L). Flow cytometric analysis further confirmed that the numbers of MNPs and, more specifically, CX3C motif chemokine receptor 1–positive (CX3CR1$^+$) microglia, were comparable between RD-RD mice and HFD-RD mice (fig. S1, D to H). Thus, a past history of obesity leads to persistent changes in later life that enhance neovascularization in retinal tissue without affecting the absolute numbers of lesion-associated MNPs.

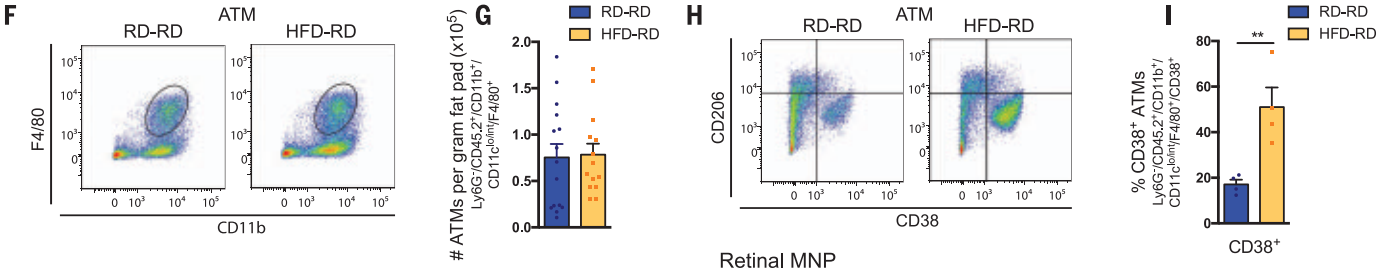
Given that adipose tissue undergoes morphological and functional changes during obesity and is one of the largest immunologically active organs, we tested whether adipose tissue from formerly obese mice displayed any residual properties that might contribute to disease. We conducted adipose tissue transplantations (ATTs) in which equal amounts (500 mg) of epididymal visceral white adipose tissue (eWAT) fat pads from RD-RD mice or HFD-RD mice were transplanted into C57BL/6J recipient mice (*28*, *29*) (Fig. 1M). All mice progressively gained weight from 1 week after ATT, and no significant differences in weight were observed among groups (fig. S2A). Surgery-induced inflammation had subsided in recipient mice within 7 days, as demonstrated by the normalization of plasma tumor necrosis factor (TNF), as well as circulating MNPs and neutrophils (fig. S2, B to I). To verify successful ATT, we confirmed that grafts were healthy and vascularized without necrosis (Fig. 1N). Three weeks after transplantation, recipient mice were subjected to laser-induced CNV. Mice receiving HFD-RD ATT developed more CNV than RD-RD ATT recipient mice or sham-operated mice (Fig. 1, O to R). The number of MNPs recruited to sites of injury also did not differ significantly (P = 0.55) between groups (Fig. 1S). Collectively, these data demonstrate that despite normalization of body weight and correction of metabolic abnormalities, adipose tissue from formerly obese mice retains properties that promote pathological neovascularization after experimental injury in distal tissues.

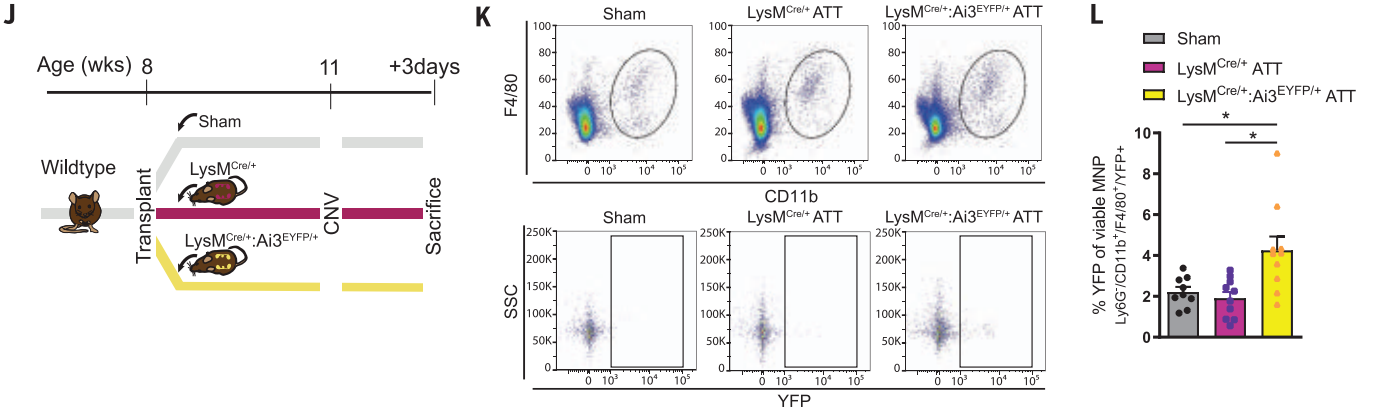
Adipose tissue macrophages are primed for cytokine production by past obesity and maintain proinflammatory profiles after weight loss

To investigate the source of angiogenic memory in the visceral adipose tissue of HFD-RD mice, we first conducted a gross histological assessment with hematoxylin & eosin (H&E) staining (Fig. 2A). The weight of adipose tissue in HFD-RD mice returned to levels similar to that in RD-RD mice, whereas that of mice fed


A
HFD
RD-RD
HFD-RD
50µm
50µm
50µm


B
Fat pad weight per body weight (g/g)
HFD
RD-RD
HFD-RD
C
AUC of % total adipocytes
Large adipocytes (> 7000 µm²)
D
AUC of % total adipocytes
Small adipocytes (≤ 7000 µm²)
E
Crown-like structures (cells/mm²)


F
ATM
RD-RD
HFD-RD
F4/80
CD11b
G
ATMs per gram fat pad (x10⁵)
Ly6G⁻/CD45.2⁺/CD11b⁺/CD11c^lo/int/F4/80⁺
H
CD206
CD38
I
% CD38⁺ ATMs
CD38⁺


J
Age (wks) 8 11 +3days
Sham
Wildtype
Transplant
LysM^Cre/+
LysM^Cre/+:Ai3^EYFP/+
CNV
Sacrifice
K
Retinal MNP
Sham
LysM^Cre/+ ATT
LysM^Cre/+:Ai3^EYFP/+ ATT
F4/80
CD11b
SSC
YFP
L
% YFP of viable MNP
Ly6G⁻/CD11b⁺/F4/80⁺/YFP+

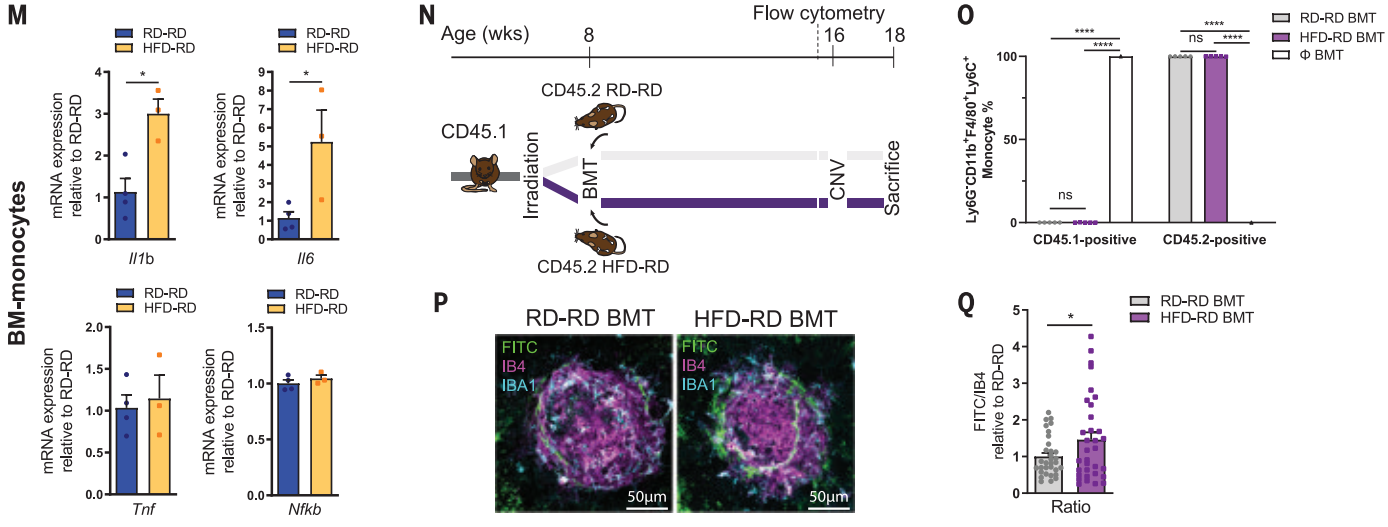

M
BM-monocytes
mRNA expression relative to RD-RD
Il1b
Il6
Tnf
Nfkb
N
Flow cytometry
Age (wks) 8 16 18
CD45.2 RD-RD
CD45.1
Irradiation
BMT
CNV
Sacrifice
CD45.2 HFD-RD
O
Ly6G⁻CD11b⁺F4/80⁺Ly6C⁺ Monocyte %
RD-RD BMT
HFD-RD BMT
Φ BMT
CD45.1-positive
CD45.2-positive
P
RD-RD BMT
HFD-RD BMT
FITC
IB4
IBA1
50µm
Q
FITC/IB4 relative to RD-RD
Ratio

Fig. 2. ATMs are primed by prior obesity and maintain a proinflammatory profile after weight loss. (**A**) H&E staining of eWAT sections from HFD mice (HFD feeding for 11 weeks), RD-RD mice, and HFD-RD mice. Scale bar, 50 μm. (**B**) Fat pad weight per body weight (g/g) from eWAT of HFD (n = 8), RD-RD (n = 6), and HFD-RD (n = 6) mice. (**C** and **D**) Adipocyte sizes in eWAT sections of HFD, RD-RD, and HFD-RD mice (n = 6). Area under the curve (AUC) of the percent total of large [> 7000 μm^2 (C)] and small [≤ 7000 μm^2 (D)] adipocytes. (**E**) Quantification of crown-like structures in adipose tissue sections of HFD (n = 6), RD-RD (n = 5), and HFD-RD (n = 6) mice. (**F** to **I**) Flow cytometry analysis of eWAT from RD-RD and HFD-RD mice. (F) Representative flow plots of ATMs in eWAT of RD-RD and HFD-RD mice. (G) Quantification of the number of macrophages ($Ly6G^-$/$CD45.2^+$/$CD11b^+$/$CD11c^{lo/int}$/ $F4/80^+$) per gram fat pad in eWAT of RD-RD (n = 15) and HFD-RD (n = 14) mice. (H) Representative flow plots of $CD38^+$ and $CD206^+$ ATMs in eWAT of RD-RD and HFD-RD mice. (I) Quantification of $CD38^+$ macrophages ($Ly6G^-$/$CD45.2^+$/$CD11b^+$/$CD11c^{lo/int}$/ $F4/80^+$/ $CD38^+$) in RD-RD mice and HFD-RD mice (n = 4). (**J**) Experimental schematic of ATT in which recipient mice were transplanted with eWAT fat pads at 8 weeks of age from either $LysM^{Cre/+}$:$Ai3^{EYFP/+}$ or $LysM^{Cre/+}$ donor mice. Sham surgeries of controls were performed similarly but without ATT. Mice were subjected to laser burns at 11 weeks and euthanized 3 days later. (**K**) Representative flow plots of MNPs ($Ly6G^-$/$CD11b^+$/$F4/80^+$/$CX3CR1^+$/$CD45.2^+$) and YFP^+ MNPs in retinas from sham-operated and $LysM^{Cre/+}$ ATT and $LysM^{Cre/+}$:$Ai3^{EYFP/+}$ ATT recipient mice. (**L**) Percentage of YFP^+ cells of viable MNPs in sham (n = 9) and $LysM^{Cre/+}$ ATT (n = 10) and $LysM^{Cre/+}$:$Ai3^{EYFP/+}$ ATT (n = 10) recipient mice. (**M**) mRNA expression of *Il1b*, *Il6*, *Tnf*, and *Nfκb* in monocytes isolated from BM of mice 72 hours after laser burn relative to RD-RD; n = 4 RD-RD, n = 3 HFD-RD. (**N**) Experimental schematic of BMT in which lethally irradiated B6.SJL (CD45.1) recipient mice were reconstituted with BM cells at 8 weeks of age from RD-RD or HFD-RD C57BL/6J (CD45.2) mice. Blood was collected from recipient mice for flow cytometry analysis 8 weeks after BMT, and then recipient mice were subjected to laser-induced CNV. (**O**) Quantification of $CD45.1^+$ and $CD45.2^+$ circulating monocytes ($Ly6G^-$/$CD11b^+$/ $F4/80^+$/$Ly6C^+$) of a $CD45.1^+$ control mouse and an irradiated $CD45.1^+$ mouse having received $CD45.2^+$ RD-RD or HFD-RD BM. (**P**) Representative confocal images showing IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in RD-RD and HFD-RD BMT mice. Scale bar, 50 μm. (**Q**) Quantification of the ratio of FITC/IB4 per laser burn relative to RD-RD mice; n = 32 burns for RD-RD BMT, n = 33 burns for HFD-RD BMT. Data information: Comparisons between groups were analyzed using one-way ANOVA with Tukey's multiple-comparisons test [(B) to (E), (L), and (O)] or Student's unpaired t test [(G), (I), (M), and (Q)]. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Error bars represent mean ± SEM.

HFD throughout the course of experimentation (HFD mice) remained significantly greater (Fig. 2B). The mean of frequency of adipocyte size did not differ between RD-RD and HFD-RD mice (fig. S3, A and B); we however found a similar increase in the number of large adipocytes (>7000 μm) in HFD and HFD-RD mice compared to RD-RD mice and a larger number of small adipocytes (≤ 7000 μm) in RD-RD mice (Fig. 2, C and D). These results suggest that obesity-induced changes in the size and distribution of adipocytes persist long after weight loss.

Several immune cell subsets in adipose tissue have important roles in modulating obesity-associated inflammation (*30*–*32*). Although prior exposure to HFD feeding did not grossly affect T cell numbers (fig. S3, C to G) or T cell–mediated cytokine production (fig. S3, H to J), a substantial difference in adipose tissue macrophage (ATM) phenotype was observed. Crown-like structures, which form a syncytium of macrophages surrounding dead adipocytes in obese mice (*33*), remained a prominent feature of visceral adipose tissue in HFD-RD mice after weight loss (Fig. 2, A and E). Further, although the number of ATMs per gram of fat pad was similar between RD-RD mice and HFD-RD mice (Fig. 2, F and G, and fig. S3, C and K), flow cytometric analysis of the tissue revealed a threefold increase in proinflammatory ATMs [$F4/80^+$ or EGF-like module-containing mucin-like hormone receptor-like 1 (EMR1), $CD11b^+$, $CD38^+$] (*34*) in HFD-RD mice compared with RD-RD mice 3 days after laser burn (Fig. 2, H and I, and fig. S3L). Even before laser injury, a greater number of ATMs were $CD38^+$ in both HFD mice and HFD-RD mice than in RD-RD controls (fig. S3, M and N).

To further investigate how ATMs influence CNV, we performed ATT with fluorescently labeled ATMs. We used lysozyme 2 $(LysM)^{Cre/+}$ mice, which express Cre in the myeloid cell lineage, and bred them with $Ai3^{EYFP/+}$ mice, which harbor a targeted mutation of the *Gt(ROSA)26Sor* (R26) locus with the *loxP*-flanked STOP cassette, preventing the transcription of enhanced yellow fluorescent protein (EYFP). The resulting mice had fluorescent monocytes, mature macrophages, and granulocytes. We transplanted eWAT fat pads from $LysM^{Cre/+}$:$Ai3^{EYFP/+}$ mice or $LysM^{Cre/+}$ mice into C57BL/6J recipient mice (Fig. 2J). CNV was induced 3 weeks after transplantation, and once surgical inflammation had subsided (fig. S2, B to I), we assessed whether myeloid cells within transplanted fat pads could migrate and contribute to AMD pathology locally. Three days after CNV induction, we detected $EYFP^+$ MNPs in retinas of mice that received eWAT fat pads from $LysM^{Cre/+}$:$Ai3^{EYFP/+}$ donor mice. Transplants from $LysM^{Cre/+}$ mice were used as controls to determine background autofluorescence. Thus, it appears possible that myeloid cells from adipose tissue directly infiltrate CNV lesions and locally contribute to disease progression (Fig. 2, K and L).

Bone marrow transfer from mice with past obesity aggravates CNV in lean mice

The myeloid compartment of the bone marrow (BM) is affected by obesity, so it is possible that BM-derived myeloid cells from HFD mice might also affect CNV. We observed a significantly higher frequency of myeloid-biased multipotent progenitor cells (MPP3) in HFD mice, a trend that was similar in HFD-RD mice compared with RD-RD controls (fig. S4, A to D). We assessed transcript levels of innate immunity–related genes in monocytes isolated from the BM of RD-RD mice and HFD-RD mice using real-time quantitative PCR (RT-qPCR) and found that levels of interleukin 1β (*Il1b*) and *Il6* rose significantly in HFD-RD mice compared with RD-RD controls (Fig. 2M). To test whether the BM from formerly obese mice might contribute to the development of CNV, we performed BM transfers (BMTs) from RD-RD or HFD-RD C57BL/6J (CD45.2) mice into lethally irradiated B6.SJL (CD45.1) recipient mice (Fig. 2N and fig. S4, E and F). Eight weeks after BMT, the circulating population of monocytes expressed the donor mouse allelic variant of the pan-hematopoietic cell marker CD45, attesting to successful transfer (Fig. 2O and fig. S4G). Most ATMs in recipient mice were derived from donor mice (fig. S4H), indicating that ATMs were replaced by BM-derived cells in irradiated recipient mice. When chimeric mice were subjected to laser-induced CNV 8 weeks after BMT, those that had received BMT from HFD-RD mice developed more CNV than recipients of RD-RD BMT (Fig. 2, P and Q, and fig. S4, I and J). This occurred despite a lack of difference in the number of MNPs recruited to the sites of injury (fig. S4K). Together, these data demonstrate that BM-derived myeloid cells transferred from formerly obese mice eventually repopulate adipose tissue and retain a phenotype that drives CNV.

Past obesity induces persistent epigenomic reprogramming of ATMs toward enhanced angiogenic and inflammatory responses

We next characterized potential epigenetic changes in ATMs induced by HFD-driven obesity. Transplanted adipose tissue or BM-derived myeloid cells retained the effects of HFD exposure and exacerbated pathological neovascularization (Figs. 1 and 2), and ATMs from HFD mice persistently remained in a proinflammatory state ($F4/80^+$, $CD11b^+$, $CD38^+$) after animals were returned to a RD (Fig. 2). We therefore

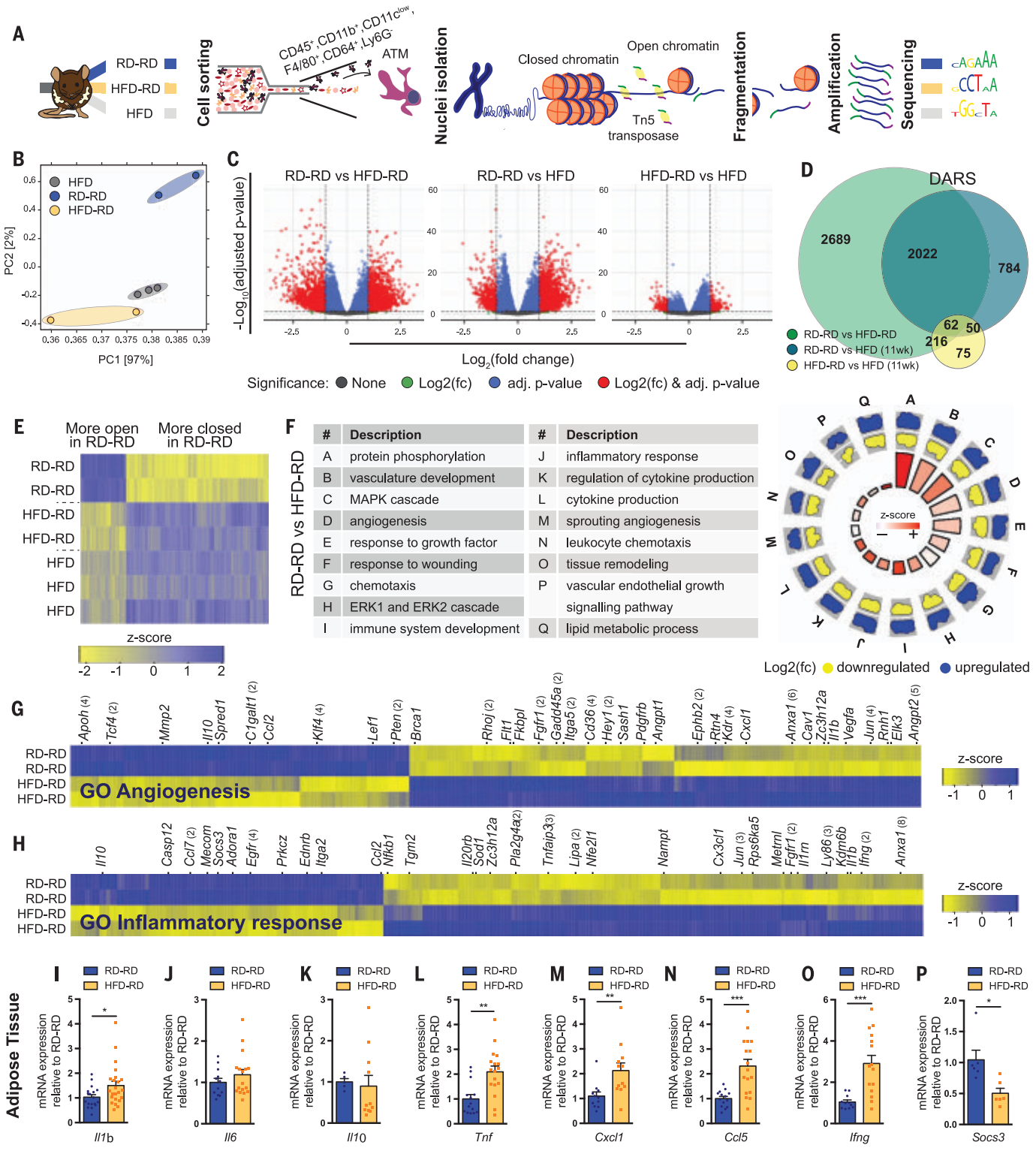


Fig. 3. Prior obesity induces epigenomic reprogramming of ATMs toward proinflammatory and proangiogenic phenotypes. (**A**) Experimental schematic of the ATAC-seq of ATMs. ATMs derived from RD-RD, HFD-RD, and HFD mice (HFD feeding for 11 weeks) were sorted by FACS and subjected to ATAC-seq. (**B**) Principal component analysis of ATAC-seq normalized read counts in peaks from HFD, RD-RD, and HFD-RD ATMs. (**C**) Volcano plots of accessible regions with DARs, defined by an adjusted $P < 0.05$ and a $|\log2(\text{fold change})| > 1.0$, between RD-RD versus HFD-RD ATMs, RD-RD versus HFD ATMs, and HFD-RD versus HFD ATMs as found by ATAC-seq. (**D**) Venn diagram showing DARs identified in the comparisons RD-RD versus HFD-RD ATMs, RD-RD versus HFD ATMs, and HFD-RD versus HFD ATMs. (**E**) Heatmap of 2022 DARs (z-score of normalized count) identified in the comparison RD-RD versus HFD ATMs, which were shared in the comparison RD-RD versus HFD-RD ATMs but not in the comparison HFD-RD versus HFD ATMs. (**F**) GO circle plot displaying gene annotation enrichment analysis in the comparison of RD-RD versus HFD-RD ATMs. The inner ring is a bar plot in which the height of the bar indicates the

significance of the term ($-\log_{10}$ adjusted *P* value) and the color corresponds to the *z*-score. The outer ring displays scatterplots of the expression levels (log fold change) for the genes in each term. (**G** and **H**) Heatmaps of DARs (z-score of normalized count) identified in the comparison of RD-RD versus HFD-RD ATMs with their nearest genes in the gene sets GO angiogenesis (G) and GO inflammatory response (H). AP-1 target genes are highlighted. If multiple DARs correspond to the same gene, then the number is indicated behind the gene. (**I** to **P**) mRNA expression in eWAT 48 hours after laser burn relative to RD-RD of *Il1b* (I), $n = 18$ for RD-RD, $n = 23$ for HFD-RD; *Il6* (J), $n = 12$ for RD-RD, $n = 17$ for HFD-RD; *Il10* (K), $n = 5$ for RD-RD, $n = 11$ for HFD-RD; *Tnf* (L), $n = 14$ for RD-RD, $n = 17$ for HFD-RD; *Cxcl1* (M), $n = 11$ for RD-RD, $n = 12$ for HFD-RD; *Ccl5* (N), $n = 14$ for RD-RD, $n = 17$ for HFD-RD; and *Ifng* (O), $n = 11$ for RD-RD, $n = 15$ for HFD-RD *Socs3* (P), $n = 6$. Data information: Comparisons between groups were analyzed using a Student's unpaired *t* test [(I) to (P)]; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$; error bars represent mean ± SEM.

evaluated global chromatin accessibility using an assay for transposase-accessible chromatin sequencing (ATAC-seq). We performed ATAC-seq on nuclei extracted from sorted ATMs from HFD, RD-RD, and HFD-RD mice before laser-induced CNV (Fig. 3A). Principal component analysis of open chromatin regions showed that samples segregated in accordance to their experimental group, with RD-RD being the most distinct and having the highest variability (Fig. 3B).

Hence, we analyzed specific and common epigenetic changes between groups. Comparison of the groups identified a total of 5898 differentially accessible regions (DARs) defined using adjusted $P < 0.05$ and $|\log2(\text{fold change})| > 1.0$ (Fig. 3, C and D). Intergroup comparisons revealed the highest diversity, and thus the greatest number of DARs between ATMs from RD-RD and HFD-RD mice (4989 DARs), and considerably fewer DARs were identified between ATMs from HFD-RD versus HFD mice (403 DARs) (Fig. 3D). A total of 2022 DARs were shared by both the RD-RD versus HFD mice, and the RD-RD versus HFD-RD mice, but not by HFD-RD versus HFD mice, indicating that these regions of chromatin in ATMs from HFD-RD mice remained in a conformation similar to the HFD state even after return to RD (Fig. 3, D and E). Among these 2022 DARs, 1535 (75.9%) were more accessible in the HFD ATMs (11 weeks) and HFD-RD ATMs, whereas 487 (24.1%) were more accessible in the RD-RD ATMs. By contrast, only 50 DARs were commonly identified in RD-RD versus HFD, and HFD-RD versus HFD mice, but not in RD-RD versus HFD-RD mice. This indicates that fewer regions of open chromatin in HFD mice reverted back to a state seen in RD once mice were returned to RD; most remained in a conformation similar to the HFD state. These data suggest that sustained changes in the chromatin accessibility landscape of ATMs from HFD mice persist long after return to RD.

Association of DARs with the nearest gene and gene ontology (GO) enrichment analysis revealed considerable pathway enrichment in the HFD-RD group compared with the RD-RD group (Fig. 3F). GO-listed genes involved in pathways coding for angiogenesis, cytokine production, and inflammatory response were enriched in the HFD-RD group. In gene sets from the GO terms "angiogenesis" and "inflammatory response," AP-1 target genes such as *Il1b*, nuclear factor of kappa light polypeptide gene enhancer in B cells 1 (*Nfkb1*), *Tnf* alpha-induced protein 3 (*Tnfaip3*), vascular endothelial growth factor A (*Vegfa*), angiopoietin 1 (*Angpt1*), platelet-derived growth factor receptor (*Pdgfr*), and beta polypeptide (*Pdgfrb*) were more accessible, whereas *Il10* was less accessible in HFD-RD ATMs (Fig. 3, G and H, and fig. S5, A and B). Together, these data demonstrate that most changes in the chromatin landscape induced by HFD feeding are maintained as open chromatin positions for at least 9 weeks, suggesting that HFD feeding leads to long-term changes in ATMs and may render them prone to proangiogenic and proinflammatory responses.

To determine whether the modifications in the chromatin landscapes described above potentiated phenotypic changes in cytokine production, we next assessed transcript abundance of innate immunity–related genes by RT-qPCR. At 48 hours after laser injury, visceral adipose tissue levels of *Il1b*, *Tnf*, chemokine (C-X-C motif) ligand 1 (*Cxcl1*), chemokine (C-C motif) ligand 5 (*Ccl5*), and interferon gamma (*Ifng*) rose significantly in HFD-RD mice compared with RD-RD controls, supporting the idea that epigenetic changes in HFD-RD mice may lead to biased expression of proinflammatory genes (Fig. 3, I to P). Consistent with this, expression of Suppressor of Cytokine Signaling-3 (*Socs3*), which is a negative regulator of cytokine signaling, was down-regulated. Furthermore, serum concentrations of proinflammatory cytokines such as IL1β, IL2, IL12, TNF-α, and IFNγ were increased in HFD-RD. Anti-inflammatory IL10 also increased (fig. S6, A to G). Collectively, these data suggest that obesity leads to persistent proinflammatory changes in visceral adipose tissue and predisposes to increased systemic inflammation after experimentally induced injury.

Stearic acid potentiates macrophage memory through activation of TLR4 signaling

During obesity, several effectors such as free fatty acids (FFAs), triglycerides, ceramides, gut-derived endotoxin, and damage-associated molecular patterns (DAMPs) such as S100A8-A9 and High Mobility Group Box1 (HMGB1) are thought to activate ATMs and other cells of the innate immune system (*26*, *35*–*38*). To elucidate potential lipids that might drive epigenetic changes caused by HFD, we used quantitative fatty acid (FA) profiling by gas chromatography–mass spectrometry (GC-MS). Stearic acid (SA) (C18:0) was the most abundant FA bound to plasma phospholipids of mice fed a HFD (time point 1, Fig. 4A and fig. S7, A and B). Additionally, we observed higher amounts of proinflammatory arachidonic acid (C20:4) also bound to phospholipids from mice fed a HFD. By contrast, we did not observe enrichment in plasma concentrations of SA or any other analyzed FA in HFD-RD mice at time point 2 after return to RD and metabolic normalization (Fig. 4B and fig. S7C), indicating that SA only accumulated in the phospholipid fraction during HFD feeding. Palmitic acid (C16:0) did not differ between the two groups in our experiments (Fig. 4A).

Saturated FFAs, including SA and palmitic acid, induce inflammation through Toll-like receptor (TLR) 4-JNK or nuclear factor κB (NF-κB) signaling (*36*, *39*). We therefore differentiated BM cells from 8-week-old male C57BL/6J mice into mature macrophages and stimulated them with SA or palmitic acid with or without TAK-242 (an inhibitor of TLR4 signaling) (fig. S8, A to F). After acute stimulation for 6 hours, SA induced the expression of proinflammatory and proangiogenic genes such as *Tnf*, *Tnfaip3*, *Il6*, chemokine (C-C motif) ligand 2 (*Ccl2*), *Cxcl1*, *Nfκb*, and *Vegf* in BM-derived macrophages (BMDMs) (fig. S8B). Pretreatment with TAK-242 abolished the induction of the *Il1b*, *Tnf*, *Tnfaip3*, *Il6*, *Ccl5*, and *Cxcl1* genes (fig. S8B). Similarly, palmitic acid induced the expression of proinflammatory and proangiogenic genes in a TLR4-dependent manner (fig. S8E).

Given that SA was the most abundant phospholipid-bound FA during the obese phase of the WGWL experiments (Fig. 4A), we investigated its potential to affect macrophage memory. We treated BMDMs with SA for 24 hours and then stimulated them with lipopolysaccharide (LPS) 5 days after the removal of SA. Compared with control macrophages pretreated with bovine serum albumin (BSA), SA-pretreated macrophages had increased expression of mRNAs encoding *Tnf*, *Tnfaip3*, *Il6*, and *Nfκβ* in response to secondary stimulation (Fig. 4, C and D). We observed heightened expression of *Il1b*, *Il6*, and *Tnf* in SA-pretreated macrophages in response to secondary stimulation after a wash period of 10 days (fig. S8C). The elevated expression of *Il6* persisted for at least another 10 days, and *Il1b* and *Tnf* showed the same trends (fig. S8C). Pretreatment with TAK-242 decreased the effect of SA on macrophage memory (Fig. 4D and fig. S8C). Consistent with a role for TLR4 in this process, SA-pretreated macrophages subsequently stimulated with

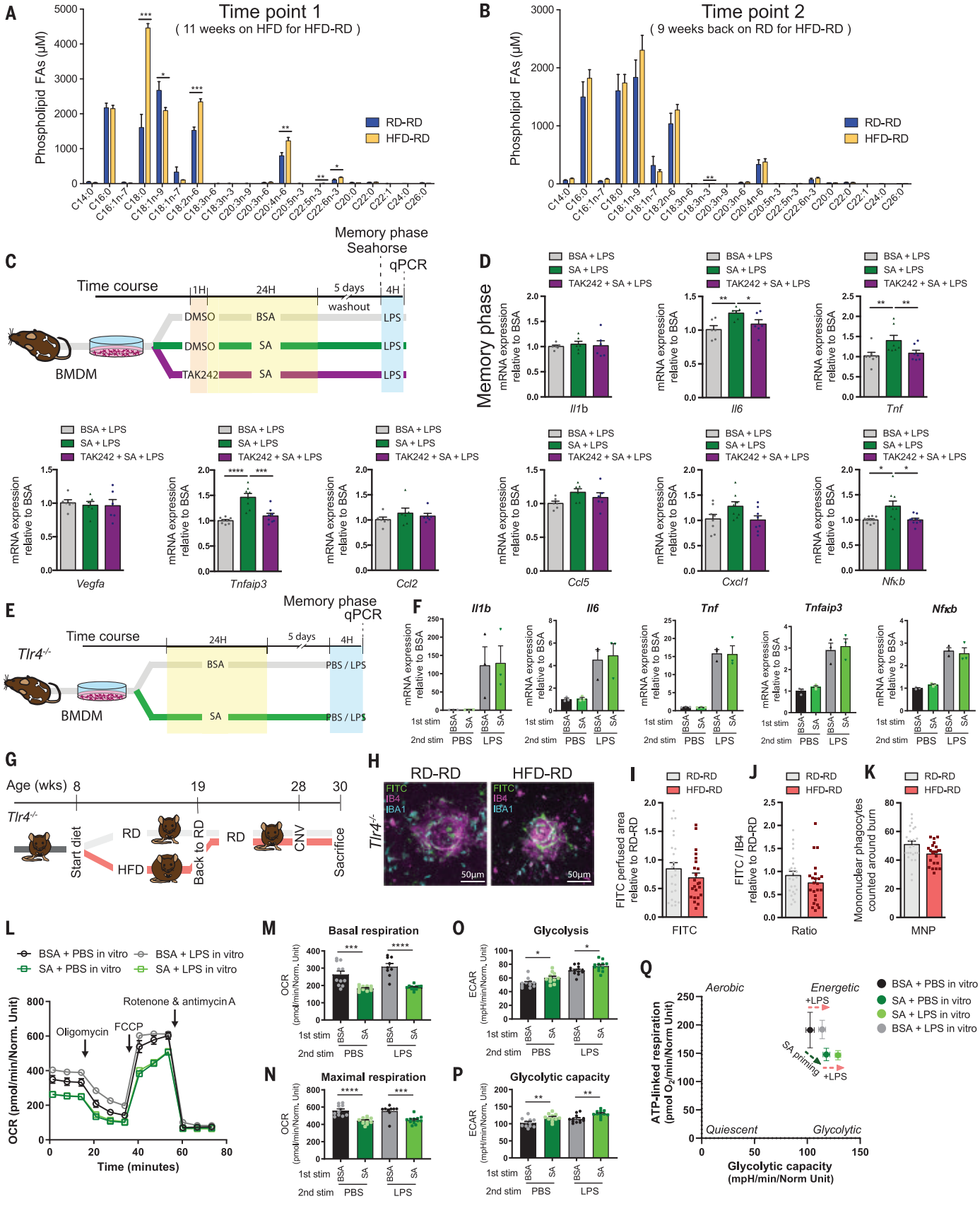


Fig. 4. SA potentiates macrophages through activation of TLR4 signaling and induces sustained metabolic rewiring. (**A** and **B**) Concentration of FAs in plasma phospholipids in RD-RD (n = 6 each time point) and HFD-RD (n = 6 each time point) mice at 19 weeks (time point 1, 11 weeks HFD

for HFD-RD mice) (A) and 28 weeks (time point 2, 9 weeks back on RD for HFD-RD mice) (B). (**C**) Experimental schematic of the in vitro SA-induced immune memory model. BMDMs from C57BL/6J mice were pretreated with DMSO (control) or a TLR4-specific inhibitor (TAK-242) for 1 hour. The BMDMs were then stimulated with BSA (control) or SA. After 24 hours of stimulation, BMDMs were cultured in basal medium for another 5 days before a secondary stimulation with LPS for 4 hours (memory phase). Total RNA was extracted and analyzed for gene expression by qPCR. Seahorse assay was performed on BMDMs without TAK-242 pretreatment before secondary stimulation with LPS [(L) to (Q)]. (**D**) BMDM mRNA expression of *Il1b*, $n = 6$; *Il6*, $n = 6$; *Tnf*, $n = 7$; *Vegfa*, $n = 6$; *Tnfaip3*, $n = 8$; *Ccl2*, $n = 6$; *Ccl5*, $n = 6$; *Cxcl1*, $n = 8$; and *Nfκb*, $n = 10$ (BSA + LPS) and $n = 9$ (SA + LPS and TAK-242 + SA + LPS) as indicated in (C). (**E**) Experimental schematic of the in vitro SA-induced immune memory model in *Tlr4* knockout ($Tlr4^{-/-}$) mice. BMDMs from $Tlr4^{-/-}$ mice were stimulated with BSA (control) or SA for 24 hours. After 5 days of washout, BMDMs were restimulated with or without LPS for 4 hours before gene expression analysis by qPCR. (**F**) BMDM mRNA expression of *Il1b*, *Il6*, *Tnf*, *Tnfaip3*, and *Nfκb* ($n = 3$ per condition) as indicated in (E). (**G**) Experimental schematic in which half of the $Tlr4^{-/-}$ mice were started on a HFD at 8 weeks of age and switched back to a RD at 19 weeks; this group is referred to as HFD-RD. Control mice were fed a RD throughout and are referred to as RD-RD. Mice were subjected to laser-induced CNV at 28 weeks and euthanized at week 30, when eyes were collected. (**H**) Representative confocal images of IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in RD-RD and HFD-RD *Tlr4* knockout mice. Scale bar, 50 μm. (**I** and **J**) Quantification of the FITC-dextran–labeled CNV area (I) and the ratio of FITC/IB4 per laser burn (J) relative to RD-RD *Tlr4* knockout mice at week 30 (day 14); $n = 24$ burns for RD-RD, $n = 23$ burns for HFD-RD. (**K**) Number of $IBA1^+$ MNPs around the laser impact area relative to RD-RD *Tlr4* knockout mice at day 14; $n = 24$ burns for RD-RD, $n = 23$ burns for HFD-RD. (**L**) OCR of BMDMs from BSA- and SA-pretreated mice with or without LPS restimulation as indicated in (C); $n = 11$ (BSA + PBS in vitro), $n = 11$ (SA + PBS in vitro), $n = 9$ (BSA + LPS in vitro), $n = 12$ (SA + LPS in vitro). (**M** and **N**) Basal respiration (M) and maximal respiration (N) of BMDMs from BSA- and SA-pretreated mice with or without LPS restimulation; $n = 11$ BSA + PBS in vitro, $n = 11$ SA + PBS in vitro, $n = 9$ BSA + LPS in vitro, and $n = 12$ SA + LPS in vitro. (**O** and **P**) Glycolysis (O) and glycolytic capacity (P) of BMDMs from BSA- and SA-pretreated mice with or without LPS restimulation; $n = 11$ BSA + PBS in vitro, $n = 11$ SA + PBS in vitro, $n = 9$ BSA + LPS in vitro, and $n = 12$ SA + LPS in vitro. (**Q**) Energy map of the four conditions tested charging ATP-linked respiration versus glycolytic capacity. Data information: Comparisons between groups were analyzed using a Student's unpaired *t* test [(A), (B), (I to K), (M), (N), (O), (P)], a multiple *t* test (F), or a one-way ANOVA with Tukey's multiple-comparisons test (D). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$; error bars represent mean ± SEM.

HMGB1 (a TLR4 ligand that provokes a lesser inflammatory response than LPS) showed greater cytokine production than control BSA-pretreated macrophages (fig. S8, G and H). By contrast, palmitic acid–pretreated macrophages showed a lower-magnitude regulation of gene expression in response to subsequent stimulation with LPS (fig. S8F).

Consistent with the results of pharmacologic inhibition of TLR4, SA-induced effects on cytokine and NF-κB mRNA were abolished in BMDMs derived from mice with a coding sequence deletion of *Tlr4* causing loss of function ($Tlr4^{Lps\text{-}del}$; hereafter referred to as $Tlr4^{-/-}$) (Fig. 4, E and F). $Tlr4^{-/-}$ BMDMs retained response to LPS or SA, likely through other TLRs (fig. S9, A and B). $Tlr4^{-/-}$ mice fed RD-RD or HFD-RD showed the same body weight change as observed in WT C57BL/6J mice (fig. S1A; Fig. 4G; and fig. S9, C and D). The systemic metabolic consequences (as shown by GTT and ITT) of DIO in $Tlr4^{-/-}$ mice were also similar to those in C57BL/6J mice (fig. S9, E to H). However, $Tlr4^{-/-}$ mice did not retain a memory phenotype, as described above (Figs. 1 and 2), and had similar magnitudes of CNV- and FITC-perfused areas as RD-RD and HFD-RD mice (Fig. 4, G to K, and fig. S9I). Thus, SA, the most enriched FA in plasma phospholipids of HFD mice, potentiates macrophages for future cytokine production through TLR4 and leads to innate memory in DIO that aggravates pathological angiogenesis in response to experimental injury.

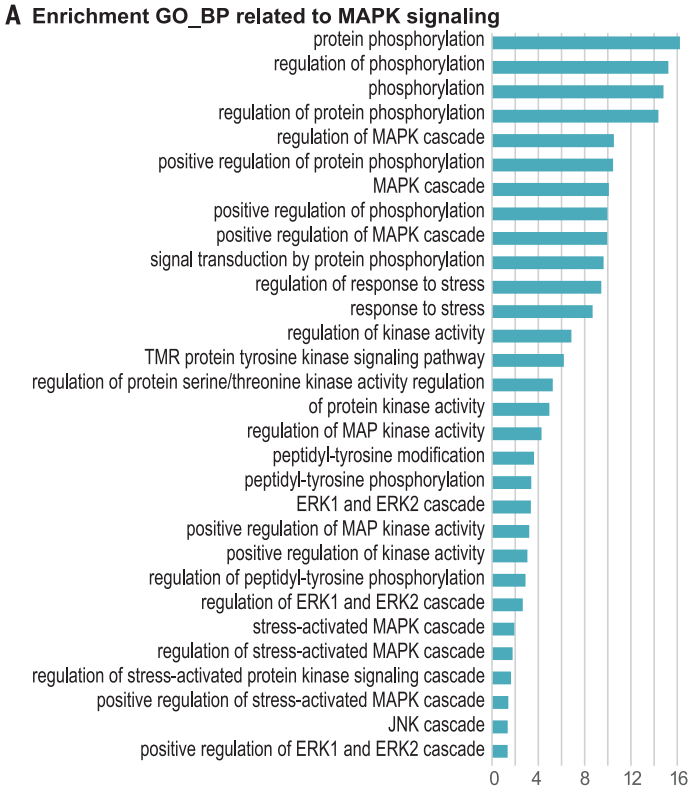
Prior exposure to SA shifts macrophage metabolism toward glycolysis

Given that cellular energy metabolism affects myeloid cell phenotype and function (*40*), we investigated whether exposure to SA provokes innate immune memory by biasing metabolism. We assessed mitochondrial respiration by measuring the oxygen consumption rate (OCR) in BSA (control) or SA-pretreated BMDMs (at memory phase 5 days after treatment; Fig. 4C). Small but statistically significant decreases were noted in basal respiration and maximal respiration in SA-pretreated BMDMs both before and after restimulation with LPS (Fig. 4, L to N).

We also measured the extracellular acidification rate (ECAR) to assess glycolysis in BSA (control) or SA-pretreated BMDMs. Consistent with a primed state for inflammatory response (Fig. 4D), SA-treated BMDMs showed small but significant shifts in glycolysis and glycolytic capacity (Fig. 4, O and P) (*41*). Subsequent treatment with LPS caused a further increase in glycolysis and glycolytic capacity in both groups (Fig. 4, O and P). We suggest that the memory state of myeloid cells is associated with metabolic reprogramming and that exposure to SA may shift myeloid cell metabolism toward glycolysis with less reliance on oxidative metabolism (Fig. 4Q), an effect that persists long after weight loss.

Obesity-driven reprogramming of macrophages correlates with chromatin remodeling at AP-1–binding sites

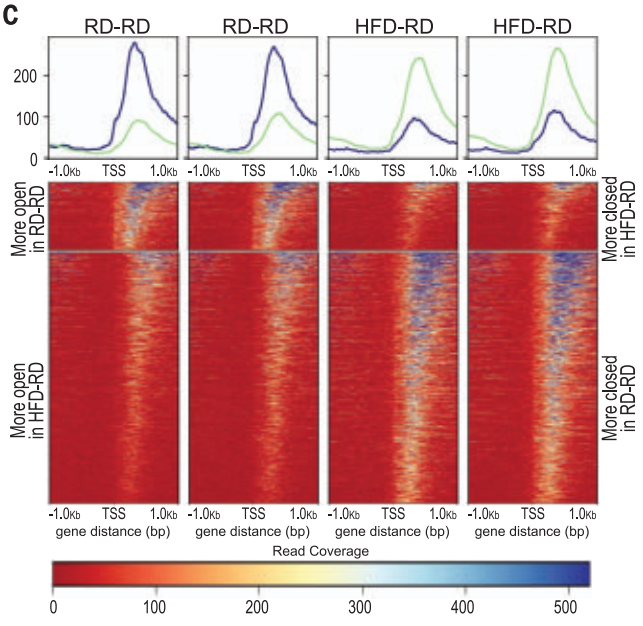
ATAC-seq from nuclei extracted from FACS-sorted ATMs of RD-RD or HFD-RD mice revealed a landscape of open chromatin regions that are enriched in proximity of genes related to the mitogen-activated protein kinase (MAPK)/c-Jun N-terminal kinase (JNK) or extracellular signal–regulated kinase (ERK) signaling pathway in HFD-RD ATMs (Fig. 5A). We thus hypothesize that genes encoding effectors of the MAPK signaling pathway might be epigenetically modulated in ATMs of formerly obese mice. Specifically, more accessible DARs in HFD-RD ATMs were enriched for regions containing consensus-binding sites for the AP-1 family of transcription factors, which comprises several members, including c-JUN, c-FOS, and ATF. The top eight highest-ranked motifs correspond to binding of AP-1 family members based on a Hypergeometric Optimization of Motif EnRichment (HOMER) motif search (Fig. 5, B and C). AP-1 binding itself leads to chromatin remodeling by recruiting histone-modifying enzymes that trigger proinflammatory genes (*42*, *43*). With this in mind, we next tested whether c-JUN, the major component of AP-1 family, and histone acetyltransferase (HAT) are recruited to the promoter of the *Tnf* gene in cells exposed to SA (Fig. 5D). A 1-hour stimulation of BMDMs with SA increased the phosphorylation of c-JUN (Fig. 5, E to G). Phosphorylation of c-JUN promotes the expression of target genes by facilitating its physical interaction with the histone acetyltransferase p300 (EP300) (*44*, *45*). Chromatin immunoprecipitation (ChIP)–qPCR assays of SA-stimulated BMDMs revealed recruitment of c-JUN to the promoter region of the *Tnf* gene in cells treated with SA (Fig. 5H). This was accompanied by recruitment of EP300 to the promoter region of the *Tnf* gene and significant increases in acetylation of histone H3 on lysine 27 (H3K27ac), leading to higher activation of transcription (*46*) (Fig. 5, I and J). However, despite a trend, statistically significant changes were not detected in the recruitment of c-JUN and EP300 to the promoter region of the *Il6* gene after stimulation with SA (Fig. 5, K to M). These data show chromatin remodeling at sites that exhibited AP-1 binding with SA as a potential mechanism of macrophage reprogramming.

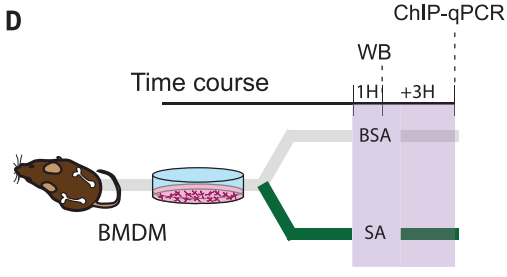


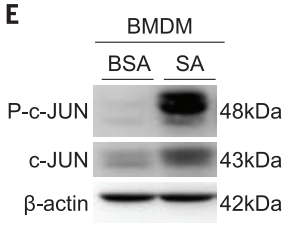
B

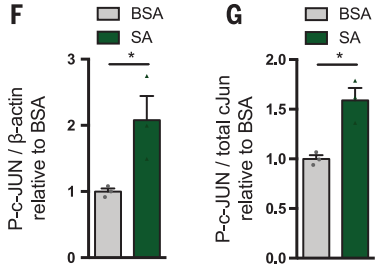
Motif	Transcription factor		P-value	Target %	Background %
	Fra1	*	1e-419	30.76	5.79
	BATF	*	1e-404	33.39	7.24
	Atf3	*	1e-402	33.01	7.11
	JunB	*	1e-391	29.89	5.86
	Fra2	*	1e-390	27.73	4.95
	AP-1	*	1e-389	34.57	8.13
	Fosl2	*	1e-355	22.49	3.39
	Jun-AP1	*	1e-309	17.93	2.35
	Bach2		1e-136	11.02	2.12
	Elf4		1e-92	22.17	9.87

* AP-1 motif









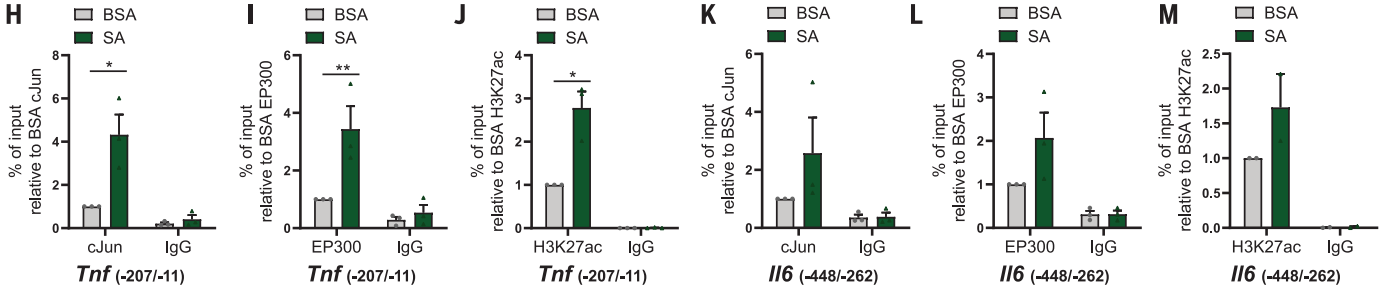


Fig. 5. Essential role of AP-1 in chromatin remodeling during obesity-driven reprogramming of macrophages. (**A**) GO enrichment analysis of MAPK signaling–related pathways enriched in ATAC-seq data from DARs between ATMs from RD-RD versus HFD-RD mice. (**B**) Top 10 enriched transcription factor

recognition sequences in ATAC-seq peaks of ATMs from the HFD-RD group on the basis of HOMER. Transcription factors indicated with asterisks are members of the AP-1 family. (**C**) Heatmaps showing density of mapped ATAC-seq reads from single biological replicates 1 kb up- and downstream of transcriptional start sites (TSS) of the AP-1 motif. (**D**) Experimental schematic in which BMDMs were stimulated with BSA or SA for 1 or 4 hours and harvested for immunoblotting or ChIP-qPCR analysis of the *Tnf* promoter, respectively. (**E**) Representative immunoblots showing c-JUN and phospho–c-JUN expression in BMDMs stimulated with BSA or SA. (**F** and **G**) Quantification of phospho–c-JUN expression in BMDMs stimulated with BSA or SA (n = 3). (**H** to **M**) ChIP-qPCR showing the enrichment of cJun [(H) and (K)], EP300 [(I) and (L)], and H3K27ac [(J) and (M)] on the chromatin of proximal promoter regions of *Tnf* and *Il6*. Data are shown as the percentage of input DNA versus IgG control antibodies. Data information: Comparisons between groups were analyzed using a Student's unpaired t test [(F) to (M)]. *P < 0.05, **P < 0.01; error bars represent mean ± SEM.

CNV is influenced by obesity-induced reprograming of ATMs and retinal myeloid cells

Given that cells of myeloid origin retain changes associated with previous obesity and that a past history of obesity influences CNV in the retina after injury, we sought to determine whether acquired innate immune memory is found in systemic myeloid cells, local retinal microglia, or both. We selectively eliminated distinct immune cell populations through diphtheria toxin (DT)–mediated apoptosis. We assessed the efficiency and specificity of two different myeloid-specific promoters, *Lysm* and *Cx3cr1*, in transgenic mice expressing fluorescent proteins. ATMs were investigated as tissue-resident cells of myeloid origin that can influence obesity-driven phenotypes. The *Lysm* promoter drove expression preferentially in ATMs compared with the *Cx3cr1* promoter (78.9 versus 30.4% of total ATMs; fig. S10, A and B). By contrast, in naïve retina and RPE-choroid-sclera complexes, the *Cx3cr1* promoter drove expression in 83.8% of retinal MNPs versus 65.9% with *Lysm*; 3 days after laser injury, this shifted to 69.4 versus 37.5%, respectively (fig. S10, C and D). *Lysm* and *Cx3cr1* promoters drove expression at similar levels in blood monocytes (69.7 versus 75.1%, respectively; fig. S10, E to G).

To deplete myeloid cells in vivo, we generated compound heterozygous mice carrying the $Lysm^{Cre}$ allele and the $R26^{iDTR}$ allele ($LysM^{Cre/+}$:$R26^{iDTR/+}$). The resulting mice expressed DT receptors specifically in myeloid cells, rendering them susceptible to targeted elimination. We investigated depletion of ATMs for proof of concept of targeted elimination of tissue-resident myeloid cells. Successful ablation of ATMs by DT was verified by flow cytometry. We observed reduction of ATMs within visceral adipose tissue in $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice (fig. S10, H and I). ATMs in control $LysM^{Cre}$ mice devoid of the $R26^{iDTR}$ allele were unaffected. No significant reduction was noted in circulating monocytes of $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice 3 days after DT injections, suggesting this population was replenished at the time of analysis (fig. S10, J and K) (*47*).

Both $LysM^{Cre/+}$:$R26^{iDTR/+}$ and control $LysM^{Cre}$ mice were put on RD-RD or HFD-RD and then subjected to laser-induced CNV (fig. S11A). Forty-eight hours after laser injury, we did not observe any changes in the serum concentration of proinflammatory cytokines between either diet paradigm in $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice (fig. S11B). Fourteen days after laser injury, CNV was more pronounced in HFD-RD $LysM^{Cre}$ mice than in RD-RD mice. This phenotype was not detected in $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice, suggesting a contribution of tissue-resident cells of myeloid origin such as ATMs in the prolonged effects of DIO (fig. S11, C to G).

To confirm whether ATMs contribute to CNV after laser injury, we transplanted eWAT fat pads from RD-RD or HFD-RD $LysM^{Cre/+}$:$R26^{iDTR/+}$ or $LysM^{Cre/+}$ mice into C57BL/6J recipient mice. Mice were then subjected to laser-induced CNV with targeted ablation of ATMs by DT injections (Fig. 6A). Fourteen days after laser injury, CNV was more pronounced in recipients of fat pads from HFD-RD $LysM^{Cre}$ control mice than in recipients of fat pads from RD-RD $LysM^{Cre}$ mice. Selective elimination of myeloid cells from adipose tissue after ATT from $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice and treatment with DT abrogated the increased CNV in recipients of fat from HFD-RD mice (Fig. 6, B to D, and fig. S11, H and I). We suggest that ATMs may retain immune memory of past obesity and contribute to experimental CNV development.

We then evaluated the contribution of local myeloid cells to the prolonged effects of obesity-induced immunity. We generated compound heterozygous mice carrying the $Cx3cr1^{CreER}$ allele and the $R26^{iDTR}$ allele ($Cx3cr1^{CreER/+}$:$R26^{iDTR/+}$). After intravitreal DT administration, ablation of local myeloid cells was verified by flow cytometry. We confirmed that MNPs, in particular $CX3CR1^{+}$ cells, were decreased in the retina and RPE-choroid-sclera complexes of $Cx3cr1^{CreER/+}$:$R26^{iDTR/+}$ mice (fig. S11, J to L). RD-RD and HFD-RD mice were intraperitoneally injected with tamoxifen for 3 consecutive days, and DT was administered intravitreally 4 and 5 weeks afterwards (Fig. 6E). Fourteen days after laser injury, $Cx3cr1^{CreER/+}$ control mice fed HFD-RD showed increased size of CNV area compared with RD-RD mice. This was in contrast to $Cx3cr1^{CreER/+}$:$R26^{iDTR/+}$ mice, in which no change was observed (Fig. 6, F to H, and fig. S11, M and N). Overall, these data reveal that myeloid cells in both tissues distal to the retina (i.e., ATMs) and local retinal myeloid cells (i.e., tissue-resident microglia and infiltrating monocytes and macrophages) appear to contribute to DIO-mediated changes in immunity. More broadly, these data suggest that a past history of obesity can reprogram local and distal effectors of the innate immune response and influence experimentally induced pathological events in CNS tissue such as the retina.

Loss of retinal function associated with light-induced retinal degeneration in previously obese mice is prevented by depletion of myeloid cells

To determine whether obesity-induced myeloid memory could influence visual function associated with AMD, we used a non-neovascular model to evaluate photoreceptor damage and retinal function in mice (*48*) (fig. S12A). Excessive light exposure in the experimental model is associated with accumulation of myeloid cells in the subretina and choroid (fig. S12, B to E) and photoreceptor damage (fig. S12, F to H), leading to retinal dysfunction (fig. S12, I to K).

After 20 weeks of diet (time point 2), RD-RD mice and HFD-RD mice were subjected to blue light-emitting diode (LED) light exposure (1500 lux for 5 days) (Fig. 6I). In vivo live imaging by optical coherence tomography (OCT) after the blue light exposure showed gradual thinning of the outer retina (photoreceptor layer) in both RD-RD and HFD-RD mice. Five days after light exposure, mice in the HFD-RD group had significantly thinner photoreceptor layers at each point of measurement (day 10) (Fig. 6, J and K, and fig. S12, L and M). Consistent with our data from the laser-induced CNV model (Fig. 1L), the number of accumulated $IBA1^{+}$ cells in the retina and choroid did not differ between the two groups (Fig. 6, L and M, and fig. S12, N to Q). Amounts of transcripts of proinflammatory genes such as *Il1b* and *Tnf* showed a small but significant rise in retina-RPE-choroid-sclera complexes in HFD-RD mice compared with control RD-RD mice (Fig. 6N). Consistent with these results, we observed significantly greater retinal dysfunction (decreased amplitude of a-wave and b-wave) in HFD-RD mice compared with RD-RD mice (Fig. 6, O to Q).

To determine whether the observed accentuated retinal dysfunction was provoked by sustained effects of past HFD on MNPs, we depleted myeloid cells in $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice. Mice were put on the RD-RD or HFD-RD regime and subjected to blue light exposure with intraperitoneal injections of DT (Fig. 6R). Retinal dysfunction was more pronounced in HFD-RD $LysM^{Cre/+}$ mice, but this phenotype was completely abrogated in $LysM^{Cre/+}$:$R26^{iDTR/+}$ mice (Fig. 6, S to U). We propose that myeloid

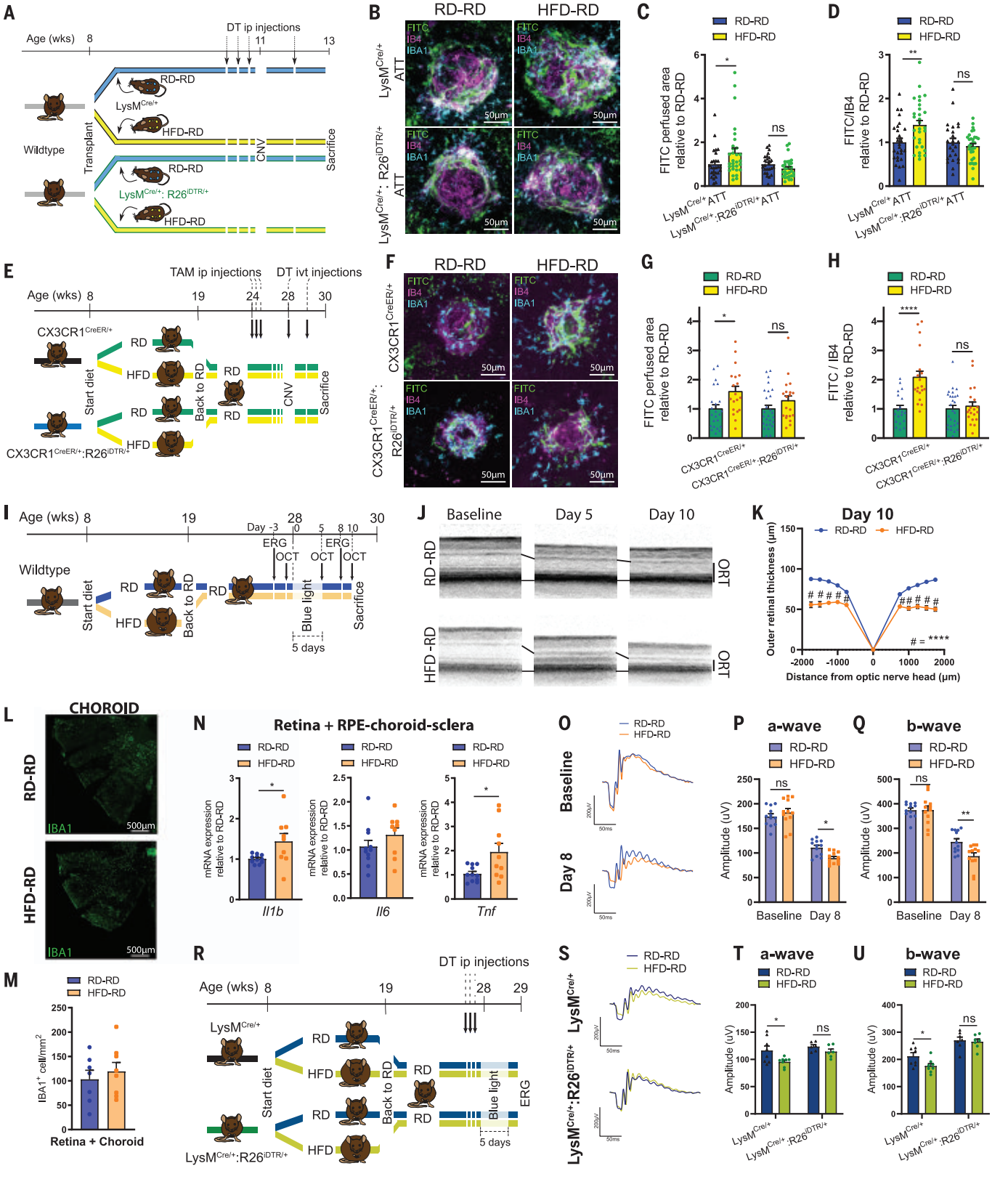


Fig. 6. Depletion of adipose tissue or retinal myeloid cells reverses a proinflammatory and proangiogenic phenotype in formerly obese mice and restores vision loss associated with retinal degeneration after light exposure. (**A**) Experimental schematic of ATT in which recipient C57BL/6J mice were transplanted with eWAT fat pads at 8 weeks of age from either LysM$^{Cre/+}$ or LysM$^{Cre/+}$:R26$^{iDTR/+}$ donor mice fed either RD-RD or HFD-RD (as shown in fig. S13A). DT was administered intraperitoneally for 3 consecutive days (at week 10), followed by one additional injection at week 12. At 11 weeks of age,

mice were subjected to laser-induced CNV and then euthanized at week 13 for analysis of eyes. (**B**) Representative confocal images showing IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in LysM$^{Cre/+}$ ATT and LysM$^{Cre/+}$:R26$^{iDTR/+}$ ATT mice after either RD-RD or HFD-RD. Scale bar, 50 μm. (**C** and **D**) Quantification of the FITC-dextran–labeled CNV area (C) and the ratio of FITC/IB4 per laser burn (D) relative to RD-RD mice at day 14; n = 27 burns for RD-RD LysM$^{Cre/+}$, n = 30 burns for HFD-RD LysM$^{Cre/+}$, n = 25 burns for RD-RD LysM$^{Cre/+}$:R26$^{iDTR/+}$, and n = 31 burns for HFD-RD LysM$^{Cre/+}$:R26$^{iDTR/+}$. (**E**) Experimental schematic in which half of the *Cx3cr1*$^{CreER/+}$ and *Cx3cr1*$^{CreER/+}$:R26$^{iDTR/+}$ mice started a HFD at 8 weeks of age and were then switched back to a RD at 19 weeks. This experimental group is referred to as HFD-RD. Control mice were fed a RD throughout and are referred to as RD-RD. Tamoxifen (TAM) was administered intraperitoneally for 3 consecutive days starting at week 24, and DT was administered intravitreally at weeks 28 and 29. At 28 weeks of age, mice were subjected to laser-induced CNV and were then euthanized at week 30 and eyes collected. (**F**) Representative confocal images showing IB4-stained laser burns with FITC-dextran–labeled CNV and IBA1-stained MNPs in *Cx3cr1*$^{CreER/+}$ and *Cx3cr1*$^{CreER/+}$:R26$^{iDTR/+}$ mice after either RD-RD or HFD-RD feeding. Scale bar, 50 μm. (**G** and **H**) Quantification of area of FITC-dextran–labeled CNV (G) and the ratio of FITC/IB4 per laser burn (H) relative to RD-RD *Cx3cr1*$^{CreER/+}$ mice at day 14; n = 23 burns for RD-RD *Cx3cr1*$^{CreER/+}$, n = 19 burns for HFD-RD *Cx3cr1*$^{CreER/+}$, n = 27 burns for RD-RD *Cx3cr1*$^{CreER/+}$:R26$^{iDTR/+}$, and n = 23 burns for HFD-RD *Cx3cr1*$^{CreER/+}$:R26$^{iDTR/+}$. (**I**) Experimental schematic of the blue LED light exposure AMD model. After 20 weeks of diet either RD-RD or HFD-RD, C57BL/6J mice were subjected to blue LED light exposure (1500 lux for 5 days) after dark adaptation overnight at 28 weeks of age. ERG was performed 3 days before (day −3) and 3 days after (day 8) light exposure. SD-OCT was performed 1 day before light exposure (baseline), just after light exposure (day 5), and 5 days after light exposure (day 10). Mice were euthanized and eyes were collected for flatmount analysis 5 days after light exposure (day 10). (**J**) Representative SD-OCT images of RD-RD and HFD-RD mice before (baseline) and after blue LED light exposure (days 5 and 10). (**K**) Outer retinal thickness (ORT) at increasing distances (−1750 μm: nasal, +1750 mm: temporal) from the center of the optic disk (0 μm) in 250-μm steps in RD-RD and HFD-RD mice 5 days after exposure (day 10); n = 10 in each group. (**L**) Representative images of IBA1-stained choroidal flatmounts at day 10. Scale bar, 500 μm. (**M**) Quantification of subretinal and choroidal IBA1$^+$ cells/mm^2 of RD-RD (n = 7) and HFD-RD (n = 8) mice. (**N**) mRNA expression in retina-RPE-choroid-sclera complexes at day 10 relative to RD-RD of *Il1b*, n = 11 for RD-RD and n = 9 for HFD-RD; *Il6*, n = 11 for RD-RD and n = 10 for HFD-RD; and *Tnf*, n = 11 for RD-RD and n = 10 for HFD-RD. (**O** to **Q**) Scotopic electroretinography before (baseline) and 3 days after light exposure (day 8). Representative images are shown in (O). (P and Q) Amplitudes of the a-waves (P) and b-waves (Q) in RD-RD and HFD-RD mice; n = 12 per each group. (**R**) Experimental schematic of LysM$^{Cre/+}$ and LysM$^{Cre/+}$:R26$^{iDTR/+}$ mice fed either RD-RD or HFD-RD and subjected to blue LED light exposure (1500 lux for 5 days) at 28 weeks of age. DT was administered intraperitoneally for 3 consecutive days at week 27. ERG was performed 3 days after (day 8) light exposure. (**S** to **U**) Scotopic electroretinography of LysM$^{Cre/+}$ or LysM$^{Cre/+}$:R26$^{iDTR/+}$ mice after receiving either RD-RD or HFD-RD. Representative images are shown in (S). Amplitudes of the a-waves (T) and b-waves (U); n = 8 for RD-RD LysM$^{Cre/+}$, n = 7 for HFD-RD LysM$^{Cre/+}$, n = 8 for RD-RD LysM$^{Cre/+}$:R26$^{iDTR/+}$, and n = 7 for HFD-RD LysM$^{Cre/+}$:R26$^{iDTR/+}$). Data information: Comparisons between groups were analyzed using a Student's unpaired t test [(C), (D), (G), (H), (M), (N), (P), (Q), (T), and (U)] or a multiple t test (K); $^*P < 0.05$, $^{**}P < 0.01$, $^{****}P < 0.0001$; error bars represent mean ± SEM.

cells retain the effects of prior DIO and can exacerbate the neuroinflammation and retinal damage associated with vision loss in models of AMD.

DISCUSSION

Although polymorphisms in several genes related to the immune and inflammatory pathways are associated with the progression of AMD (*49*), genetics can only partially explain disease incidence. We propose that in obesity, sustained systemic exposure to increased concentrations of lipids such as SA may lead to long-term changes in innate immunity and render distinct cells of myeloid lineage, such as ATMs and retinal microglia, susceptible to triggering heightened proangiogenic and proinflammatory responses. These cells may thus be primed to exacerbate CNV, even once normal weight and metabolic status have been reestablished. Targeted elimination of either cell population (ATM or retinal microglia) attenuated experimentally induced CNV in previously obese mice, consistent with a contribution of both local and peripheral reprogrammed myeloid cells to neovascular AMD.

The concept of immune memory, conventionally thought to occur uniquely with adaptive immunity, has been extended to include innate immunity with the demonstration that monocytes and macrophages undergo stable epigenetic and metabolic reprogramming (*50*). Consistent with our study, recent work suggests that Western diets trigger sterile inflammation and induce innate immune memory in the BM (*51*). Although we found absolute numbers of immune cells unaffected in formerly obese mice, we saw sustained proinflammatory phenotypes in ATMs. ATAC-seq analysis revealed preferential chromatin decondensation at binding sites for AP-1, with enrichment in accessibility regions containing consensus binding sites for several AP-1 family members such as c-JUN, c-FOS, and ATF. AP-1 binding is reported to lead to chromatin remodeling by recruitment of histone modification enzyme to a promoter of proinflammatory genes (*43*). In agreement with this, we observed HAT EP300 bound to the promoter region of the *Tnf* gene in SA-stimulated macrophages. Moreover, histone acetylation at lysine 27 of histone 3 (H3K27ac) was enriched in the promoter region of the *Tnf* gene, and AP-1 target genes such as *Vegfa*, *Il1b*, *Tnfaip2*, and *Cxcl1* were in the open conformation, which is consistent with epigenetic modification leading to a primed innate immune system.

The observed changes in the innate immune cells of formerly obese mice can be recapitulated with exposure to selected FAs. Elevated circulating saturated FAs can provoke inflammation in obesity after HFD (*52*). SA and palmitic acid activate TLR4 signaling pathways and induce the production of inflammatory mediators such as TNF-α, IL6, NF-κB, and COX2 in macrophages (*36*, *39*, *53*, *54*). We found that both of these saturated FAs increased the expression of proinflammatory and proangiogenic genes in a TLR4-dependent manner; however, only SA led to a persistent hyperresponsiveness state in macrophages. LPS (the microbial ligand for TLR4) also has the ability to trigger a long-lasting hyperresponsive state in myeloid cells (*55*), suggesting a conserved mechanism through TLR4 for initiating innate memory. Persistent epigenetic reprogramming of long-lived cells of myeloid lineage by past events associated with lipid overload identifies potential nodes that can be targeted to mitigate neuroinflammation in later life.

In summary, we show in mice that a past history of obesity has the propensity to induce long-term chromatin remodeling in tissue-resident myeloid cells such as ATMs and that this ultimately influences neuroinflammation after experimental injury in distal tissues such as the retina. These sustained changes in the chromatin landscape of innate immune cells exacerbate pathological angiogenesis and neuronal degeneration in experimental models, suggesting a link between systemic innate immune training and retinal disease. In addition, future therapeutic avenues to influence epigenetic reprogramming of the innate immune system or to eliminate subpopulations of reprogrammed innate immune cells may delay or prevent the onset of both neovascular AMD and non-neovascular AMD.

MATERIALS AND METHODS

Reagents

See table S3 for a detailed description of all reagents.

Animal studies

Mice were housed in the Hospital Maisonneuve-Rosemont Research Center animal facility in a 12-hour light-dark cycle and with ad libitum access to food and water unless otherwise indicated. All animals used in this study were males. The experimental procedures were initiated when mice were 8 weeks of age unless otherwise indicated.

C57BL/6J and B6.B10ScN-*Tlr4*$^{lps-del}$/JthJ (referred to as *Tlr4*$^{-/-}$) mice were purchased from The Jackson Laboratory (Bar Harbor, ME, USA) and bred at the Hospital Maisonneuve-Rosemont Research Center animal facility.

Homozygous B6.129P2(C)-*Cx3cr1*$^{tm2.1(cre/ERT2)Jung}$/J (referred to as *Cx3cr1*^{CreER}) mice and homozygous B6.129P2-*Lyz2*$^{tm1(cre)Ifo}$/*J* (referred to as LysM^{Cre}) mice were crossed in-house with homozygous C57BL/6-*Gt(ROSA)26Sor*$^{tm1(HBEGF)Awai}$/J (referred to as R26^{iDTR}) mice to obtain heterozygous *Cx3cr1*$^{CreER/+}$:R26 $^{iDTR/+}$ and LysM$^{Cre/+}$:R26 $^{iDTR/+}$ mice, respectively. Homozygous *B6.129P-Cx3cr1*^{tm1Litt}/*J* (referred to as *Cx3cr1*^{GFP}) mice were crossed in-house with C57BL/6J mice to obtain heterozygous *Cx3cr1*$^{GFP/+}$ mice. Homozygous B6.Cg-*Gt(ROSA)26Sor*$^{tm3(CAG-EYFP)Hze}$/J (referred to as Ai3^{EYFP}) mice were crossed in-house with homozygous LysM^{Cre} mice to obtain heterozygus LysM$^{Cre/+}$:Ai3$^{EYFP/+}$ mice.

All animal studies were performed in compliance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines and the Association for Research in Vision and Ophthalmology (ARVO) Statement for the Use of Animals in Ophthalmic and Vision Research and were approved by the Animal Care Committee of the Hospital Maisonneuve-Rosemont Research Center in agreement with the guidelines established by the Canadian Council on Animal Care.

Myeloid cell depletion

Microglia depletion was performed using *Cx3cr1*$^{CreER/+}$:R26 $^{iDTR/+}$ or LysM$^{Cre/+}$:R26 $^{iDTR/+}$ mice. In *Cx3cr1*$^{CreER/+}$:R26 $^{iDTR/+}$ mice, the activation of Cre recombinase (under the control of the *Cx3cr1* promoter) can be induced by tamoxifen treatment and leads to surface expression of DTR on CX3CR1-expressing cells. At 24 weeks of age, mice were subjected to daily intraperitoneal injections with tamoxifen diluted in corn oil (4 mg per mouse per day, stock solution at 20 mg/ml) for 3 consecutive days.

To deplete CX3CR1^{+} cells in *Cx3cr1*$^{CreER/+}$:R26 $^{iDTR/+}$ mice, DT was administered intravitreally (25 ng/μl saline per eye) at 28 and 29 weeks of age. Depletion of LysM^{+} cells in LysM$^{Cre/+}$:R26 $^{iDTR/+}$ mice was achieved by intraperitoneal injections with DT (100 ng per mouse per day) for 3 consecutive days at 27 weeks of age, followed by one additional injection at 29 weeks of age. For the blue LED light model, depletion of LysM^{+} cells in LysM$^{Cre/+}$:R26 $^{iDTR/+}$ mice was achieved by intraperitoneal injections with DT for 3 consecutive days at 27 weeks of age.

Primary BM-derived macrophages

Donor mice (8-week-old male C57BL/6J or *Tlr4*$^{-/-}$) were euthanized and leg bones were dissected. Femur and tibia cavities were flushed with phosphate-buffered saline (PBS) supplemented with 10% fetal bovine serum (FBS) using a syringe, resuspended, and passed through a 70-μm strainer. Red blood cells (RBCs) were removed using RBC lysis buffer (catalog no. 00-4333-57; eBioscience). After centrifugation, BM cells were recovered and cultivated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS and Penicilin-Streptomycin 1% (100 U/ml). Macrophage colony–stimulating factor (M-CSF) (20 ng/ml, catalog no. PMC2044; Invitrogen) was used to generate in vitro BMDMs from BM progenitor cells. After 3 days of incubation at 37°C with 5% CO_2, medium containing M-CSF was refreshed. Cells were allowed to differentiate for a total of 6 days before their medium was replaced by complete medium without M-CSF. Macrophage purity was ~99% as evaluated by flow cytometry.

METHOD DETAILS

Diet paradigm

The fat component of the HFD, lard 245 g and soybean oil USP 25 g/773.85 g, was the same as that for the RD, lard 20 g and soybean oil USP 25 g/1055.5 g. Eight-week-old mice were fed with either HFD (60% kcal fat) or RD (10% kcal fat) to study weight gain experimentally, as outlined in the text, and mouse weight was monitored regularly. The time course of the diet feeding is indicated when appropriate. Upon sacrifice, the eWAT fat pads were weighed.

Laser-induced CNV

Animals were anesthetized using a mixture of 10% ketamine and 4% xylazine (10 μl/g of body weight) injected intraperitoneally, and pupils were dilated with mydriacyl 0.5%. Using an argon laser, we ruptured their Bruch's membrane to induce CNV as described previously (*27*). Each eye received four distinct laser burns (400 mW, 100 μm, 0.05 s) distributed equidistantly and following the optic nerve head as a central reference. Disruption of the Bruch's membrane was verified through observation of a visible heat bubble at the site of injury.

CNV evaluation

Fourteen days after CNV induction, mice were anesthetized with isoflurane gas and intracardially perfused with 0.5 ml of FITC-dextran (average molecular weight, 2,000,000). FITC-dextran was left in circulation for 5 min while the animals were under anesthesia; subsequently, the mice were sacrificed and the eye globes were enucleated and fixed in paraformaldehyde (PFA) 4% at room temperature for 30 min. The RPE-choroid-sclera complex was dissected and separated from the neuroretina. Tissue was incubated for 1 hour in a blocking solution (3% BSA + 0.3% Triton X-100), followed by an overnight incubation with antibodies. Blood vessels were investigated using rhodamine-labeled *Griffonia (bandeiraea) Simplicifolia* Isolectin I (1:100), and MNPs were detected with IBA1 (1:350). After several washes with PBS, RPE-choroid-sclera complexes were incubated with secondary fluorochrome-conjugated species-appropriate antibodies for 1 hour. The tissue was then mounted onto a glass slide and imaged using an Olympus FluoView FV1000 laser scanning confocal inverted microscope (Olympus Canada, Richmond Hill, ON). For analysis, the Z-stacks were compressed into one image. The area of neovascularization (FITC-dextran^{+}) and the burn area (isolectin^{+}), as well as the number of MNPs (IBA1^{+}) were quantified using ImageJ software (version 1.0; National Institutes of Health, Bethesda, MD, USA). The quantification of the number of IBA1^{+} cells was performed blinded to diet group.

Blue LED light exposure AMD model

Mice were dark adapted overnight, and then pupils were dilated using atropine sulfate ophthalmic ointment before exposure to light. Mice were exposed to blue light from an LED (wavelength 450 nm) at a light intensity of 1500 lux for 5 days and then returned to regular conditions under a standard 12-hour light-dark cycle until sacrifice on day 5 after illumination.

Spectral domain–OCT image acquisition and measurement of retinal thickness

Spectral domain–OCT (SD-OCT) examinations using Multiline OCT (based on a Spectralis OCT, Heidelberg Engineering) were performed in blue LED light exposure AMD model at baseline, just after exposure (day 5), and 5 days after exposure (day 10). Volume scans were performed with the optic nerve head centered, and the horizontal scan images, the center of which were adjusted to the center of the optic nerve head, were used for analysis. The software used for measurement of retinal thickness was based on the built-in Spectralis OCT and provided by Heidelberg Engineering software to facilitate manual assessment of the B-scan images. The distance between the inner border of the outer plexiform layer and the inner border of the retinal pigment epithelium was calculated as outer retinal thickness (ORT). ORT and total retinal thickness were assessed at distances of 750, 1000, 1250, 1500, and 1750 μm from the optic nerve head (nasally and temporally).

MNP quantification in choroidal and retinal flatmounts of mice exposed to blue LED light

Eyes were enucleated and fixed in 4% PFA at room temperature for 60 min. Dissected

retinas and RPE-choroid-sclera complexes were incubated with anti-IBA1 (1:350) followed by secondary antibody anti-rabbit Alexa Fluor 488. Choroids and retinas were flatmounted and imaged using Zen Blue edition 3.2 software on an AxioImager Z2 microscope (Zeiss, Jena, Germany). Choroids and retina flatmounts are large and thick samples. The full images were obtained using mosaic and stitching features of the software. Z-stacks with 3-μm steps were optically sectioned using Apotome.2 with three phase images and processed using normalization, local bleaching correction, and strong Fourier filter features provided by the software. The number of MNPs ($IBA1^+$) were counted on whole RPE/choroidal flatmounts and on the outer segment side of the retina, and their density was determined using ImageJ software.

Electroretinography

The retinal function of the blue LED light exposure AMD model mice was investigated using electroretinography (ERG). ERG measurements were performed after an overnight dark adaptation. Before measurements, mice were anesthetized with 10% ketamine and 4% xylazine (10 μl/g of body weight) injected intraperitoneally. The pupils were then dilated with cyclopentolate hydrochloride 1%. Proparacaine hydrochloride 0.5% was used to anesthetize the eye. Animals were placed on a heating pad (Harvard Apparatus, Holliston, MA) during the entire recording session to maintain their body temperature at 37°C. All manipulations were performed under dim red light. ERG was recorded in the left eye by placing an electrode (DTL Plus; Diagnosys LLC) in contact with the cornea, a reference electrode was placed on the tongue, and a neutral electrode was inserted in the tail. Mice were placed under the ERG dome (Diagnosys LLC color dome model D125), and scotopic ERGs were recorded with five flashes of 25 (P)cd.s/m^2. Results were analyzed with Diagnosys version 6.63 software. The amplitude of the a-wave was measured from baseline to trough and the b-wave from the trough of the a-wave to the highest peak of the b-wave.

Serum cytokine profiling

Blood was obtained by cardiac puncture, and serum was collected after centrifugation and preserved at −80°C until analysis. The concentrations of IL1β, IL2, IL6, IL10, IL12, IFNγ, and TNF-α were measured using the Bio-Plex Pro Mouse Cytokine Th1 Panel according to the manufacturer's instructions (catalog no. L6000004C6; Bio-Rad). Cytokine levels are expressed as total pg/ml.

Plasma phospholipid and triglyceride FA profiling

Blood was collected by cardiac puncture, and plasma was obtained after centrifugation and preserved at −80°C until analysis. Quantitative profiling of FAs bound to plasma phospholipids and triglycerides was evaluated using GC-MS as described previously (*56*, *57*). In brief, total lipids were extracted with a mixture of methyl-tert-butyl ether, methanol, and water. Phospholipids and triglycerides were eluted on an aminopropyl column (Bond Elut LRC-NH2, 500 mg; Agilent Technologies). FFAs were analyzed as methyl esters after a direct transesterification with acetyl chloride/methanol on a 7890B gas chromatograph coupled to a 5977A mass selective detector (Agilent Technologies) equipped with a capillary column (J&W Select FAME CP7420; 100 × 250 m inner diameter; Agilent Technologies) and operated in the positive chemical ionization mode using ammonia as the reagent gas. Samples (2 μl) were analyzed under the following conditions: injection at 270°C in a split mode (split ratio: 50:1) using high-purity helium as the carrier gas (constant flow rate: 0.44 ml/min) and the following temperature gradient: 190°C for 25 min, increased by 1.5°C/min until 236°C. The FAs were analyzed as their $[M + NH3]^+$ ion by selective ion monitoring, and concentrations were calculated using standard curves and isotope-labeled internal standards. Total FA concentrations in plasma phospholipids and triglycerides were calculated as the sum of each FA measured. The percentage of phospholipid or triglyceride FA was calculated as the concentration of each FA over total FAs, multiplied by 100, measured in each fraction.

Transcription analysis by RT-qPCR

Mouse tissue from in vivo experiments or BM-derived macrophages from in vitro assays were harvested and snap frozen immediately upon collection. RNA extraction was performed using TRIzol reagent (catalog no. 15596026; Invitrogen) and digested with DNase I (catalog no. D4527; Sigma-Aldrich) following the manufacturer's instructions to avoid genomic DNA amplification. Reverse transcription was performed using a 5X All-In-One RT MasterMix (catalog no. G590; Applied Biological Materials Inc.), and gene expression was analyzed using Bright Green2X qPCR Master Mix-Low Rox in an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA). Primer sequences used in this study are listed in table S1. Analysis of expression was followed using the ΔΔCT method. *Actb* expression was used as the reference housekeeping gene. Statistical analysis was performed on ΔΔCT values, and the data are presented as the expression of the target genes normalized to *Actb* (fold increase).

ChIP-qPCR

ChIP-qPCR experiments were performed using the iDeal ChIP-seq kit for histones (catalog no. C01010051; Diagenode, Liège, Belgium) following the manufacturer's protocol. Specifically, 5×10^6 BM-derived macrophages were treated with 5% FA-free BSA or 200 μM SA for 6 hours. Cells were then fixed with 1% formaldehyde for 8 min to cross-link chromatin. Chromatin samples were sheared by 10 cycles of 30 s ON and 30 s OFF with the Bioruptor Pico sonicator (Diagenode), and 1% of samples derived from 1×10^6 cells were kept for input measurements. Chromatin was immunoprecipitated using antibodies to c-Jun (catalog no. 9165; Cell Signaling Technology), P300 (catalog no. MAI-16622; Invitrogen), or histone 3K27ac (H3K27ac, catalog no. ab4729; Abcam) or using 1 μg of rabbit immunoglobulin G (IgG) from the kit as a negative control. Cross-links were then reversed, and the purified DNA was amplified by qPCR using iTaq Universal SYBR Green Supermix (catalog no. 1725124; Bio Rad Laboratories) in an Applied Biosystems Real-Time PCR machine (Thermo Fisher Scientific, Waltham, MA, USA). Primers for qPCR were previously reported (*58*) and are listed in table S2. Transcription factor–binding sites in the Tnf and Il6 regulatory region between −227 and −11 base pairs (bp) and between −262 and −448 bp were selected as described previously (*59*, *60*).

Sample preparations for flow cytometric analysis

Retinas and RPE-choroid-sclera complexes were cut into small pieces and homogenized in a solution of 750 U/ml DNAse I (catalog no. D4527; Sigma-Aldrich) and 0.5 mg/ml collagenase D (catalog no. 11088866001; Roche) for 20 min at 37°C. Homogenates were filtered through a 70-μm cell strainer.

eWAT fat pads were freshly dissected, homogenized using scissors, and incubated in a solution containing 1 mg/ml collagenase type II (catalog no. C6885; Sigma-Aldrich) for 45 min at 37°C. EDTA (10 mM) was used to stop the digestion, and the samples were incubated for an additional 5 min at room temperature. The cell suspension was then filtered through a 100-μm strainer and centrifuged at 500*g* for 10 min at 4°C to separate the mature adipocytes from the stromal vascular fraction (SVF). Supernatants were discarded, and the SVF-containing pellet was resuspended with RBC lysis buffer (catalog no. 00-4333-57; eBioscience) to remove RBCs.

Blood was collected from the tail vein by bleeding after distal cut with a 21-gauge needle, and BM cells were obtained by flushing both tibias and femurs. RBC lysis buffer (catalog no. 00-4333-57; eBioscience) was used to remove RBCs.

Flow cytometry staining and analysis

Viability of the cells was assessed by staining with 7-amino-actinomycin D (20 min at room temperature) or Zombie Aqua (15 min at room temperature), and samples were incubated with LEAF-purified anti-mouse CD16/32 (catalog no. 101310; BioLegend) for 10 min at 4°C to block FC receptors. Cells were then incubated

for 25 min at 4°C with the following antibodies: for retinas and RPE-choroid-sclera complexes, BV711 anti-CD11b (catalog no. 101242; BioLegend), PE anti-CX3CR1 (catalog no. FAB5825P; R&D Systems), APC anti-CD45.2 (catalog no. 109814; BioLegend), APC/Cy7 anti-Ly-6G (catalog no. 127624; BioLegend), and PE/Cy7 anti-F4/80 (catalog no. 123114; BioLegend); for ATMs, BV711 anti-CD11b (catalog no. 101242; BioLegend), PE anti-F4/80 (catalog no. 123110; BioLegend), APC anti-CD64 (catalog no. 139305; BioLegend), FITC anti-CD38 (catalog no. 102705; BioLegend), APC/Cy7 anti-Ly-6G (catalog no. 127624; BioLegend), BV785 anti-CD11c (catalog no. 117335; BioLegend), and PE/Cy7 anti-CD206 (catalog no. 141719; BioLegend); for T cells in adipose tissue, BV711 anti-CD11b (catalog no. 101242; Biolegend), BV785 anti-CD4 (catalog no. 100551; BioLegend), PerCP/Cy5.5 anti-CD8a (catalog no. 100733; BioLegend), PE anti–TCR-β (catalog no. 109207; BioLegend), and APC/Cy7 anti-CD45R/B220 (catalog no. 103223; BioLegend); for peripheral blood after ATT, BV711 anti-CD11b (catalog no. 101242; BioLegend), FITC anti-CD3ε (catalog no. 100305; BioLegend), BV785 anti-CD45.2 (catalog no. 109839; BioLegend), PE/Cy7 anti-F4/80 (catalog no. 123113; BioLegend), BV421 anti–Ly-6G (catalog no. 127627; BioLegend), and APC anti-CD19 (catalog no. 115511; BioLegend); for peripheral blood to characterize cells of donor and host origins, BV711 anti-CD11b (catalog no. 101242; BioLegend), APC/Cy7 anti-Ly-6G (catalog no. 127624; BioLegend), PE anti-F4/80 (catalog no. 123110; BioLegend), FITC anti–Ly-6C (catalog no. 128006; BioLegend), Pacific Blue anti-CD45.1 (catalog no. 110722; BioLegend), and Alexa Fluor 700 anti-CD45.2 (catalog no. 109822; BioLegend); for BM cells, APC/Cy7 anti-CD117 (c-kit) (catalog no. 105825; Biolegend), FITC anti–Ly-6A/E (Sca-1) (catalog no. 108105; BioLegend), PE/Cy7 anti-CD34 (catalog no. 128617; BioLegend), BV421 anti-CD48 (catalog no. 103427; BioLegend), BV711 anti-CD150 (SLAM) (catalog no. 115941; BioLegend), PE anti-CD135 (Flt3) (catalog no. 135305; BioLegend), and APC mouse lineage antibody cocktail or APC mouse lineage isotype control cocktail (catalog no. 558074; BD Pharmingen).

For T cell cytokine production, cells were stimulated with phorbol myristate acetate (50 ng/ml; catalog no. P1585; Sigma-Aldrich) and ionomycin (1 μg/ml) (catalog no. I0634; Sigma-Aldrich) in the presence of BD GolgiPlug (BFA) (catalog no. 555028; BD Bioscience) for 4 hours at 37°C. Cells were permeabilized and fixed with the BD Cytofix/Cytoperm kit (catalog no. 554722; BD Bioscience) and then incubated with FITC anti-IFNγ (catalog no. 11-7311-41; eBioscience) and PE/Cy7 anti–TNF-α (catalog no. 506323; BioLegend) for intracellular staining for 25 min at 4°C.

Samples were acquired using a Fortessa X-20 cell analyzer (BD Biosciences, Mississauga, ON, Canada) and analyzed using FlowJo Software (vVersion 10.2; FlowJo, Ashland, OR, USA).

Immunohistological evaluation of adipose tissue

eWAT depots were dissected and fixed with 10% formalin overnight. Tissues were dehydrated, embedded in paraffin, and posteriorly cut in 12-μm-thick sections following standardized histological procedures. eWAT sections were then deparaffinized, rehydrated with decreasing concentrations of ethanol, and stained with Harris H&E. Sections were subsequently dehydrated and mounted with PERTEX (HistoLab Products AB) for the visualization of cellular components. Differential interference contrast images were acquired at 20× using an AxioObserver.Z1 (Live Cell Zeiss Imaging System, Zeiss, Jena, Germany). Adipocyte size was evaluated and quantified using a custom-written program in MATLAB R2019b (9.7.0.1190202, The MathWorks, Inc).

GTT

Mice were starved for 12 hours overnight. Blood glucose was measured from the tail vein using AlphaTrak2 test strips at baseline and at 15, 30, 60, 120, and 240 min after intraperitoneal injection of PBS +10% D-glucose (2 mg/kg of weight). Plasma insulin concentrations were determined by an ultrasensitive mouse insulin ELISA kit (catalog no. 90080; Crystal Chem) at each time point of the GTT.

ITT

Mice were starved for 6 hours during daylight hours. Blood glucose was measured from the tail vein using AlphaTrak2 test strips at baseline and at 30, 60, and 120 min after intraperitoneal injection of insulin (0.75 U/kg body weight).

ATT

Eight-week-old recipient mice were randomly assigned to the different groups independently of the origin of the donor eWAT. Donor mice were anesthetized, sacrificed, and their eWAT fat pads were carefully excised and weighed. The recipient mice were anesthetized with isoflurane and subjected to multiple dorsal incisions (four to six incisions) to allow subcutaneous engraftment of the same amount (~500 mg) of donor eWAT. Sham surgeries of control animals were performed in the same manner but without fat pad transplantation. Mice were closely monitored for 3 weeks before performing laser-induced CNV. Successful engraftment of the transplantation was verified by evaluating the transplanted tissue (e.g., blood vessel reperfusion and lack of necrosis) 14 days after laser-induced CNV.

Plasma TNF measurement for ATT mice

Plasma was harvested as above, and the levels of TNF-α were determined before and 2, 7, 14, and 21 days after ATT. A mouse TNF-α ELISA kit (catalog no. BMS607-3; Invitrogen) was used to detect the secretion of TNF-α in plasma according to the manufacturer's instructions.

BMT

For the generation of chimeric mice, BM cells were obtained by flushing both tibias, femurs, and iliac crests of 28-week-old RD-RD and HFD-RD C57BL/6J (CD45.2) donor mice. Eight-week-old B6.SJL (CD45.1) recipient mice were randomly assigned to the different groups independently of the origin of the donor BM cells. Recipient mice were lethally irradiated (12 Gy) and reconstituted with 5×10^6 BM cells. Mice were closely monitored, and peripheral blood chimerism was analyzed 8 weeks after reconstitution, after which laser burn–induced CNV was performed.

ATM cell sorting

ATMs were stained with the antibodies mentioned above (except for CD206 and CD38 antibodies) and sorted from the SVF of eWAT. Viable $Ly6G^-/CD45.2^+/CD11b^+/CD64^+/F4/80^+$ cells were sorted using a FACS Aria instrument (BD Biosciences, Mississauga, ON, Canada) and recovered in FBS. After sorting, ATMs were immediately prepared for ATAC-seq sample preparation.

ATAC-seq sample preparation

FACS-sorted ATMs were used for ATAC-seq, and nuclei were isolated as previously described (*61*). Briefly, isolated cells were centrifuged at 500*g* for 5 min at 4°C and then resuspended in ice-cold PBS + 0.04% BSA. Cell lysis was performed for 5 min on ice by adding 45 μl of lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20, 0.1% Nonidet P-40, 0.001% digitonin, and 1% BSA). After lysis, 50 μl of ice-cold wash buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween 20, and 1% BSA) was added and then centrifuged at 500*g* for 5 min at 4°C. After an additional wash, the nuclei were counted and used for the transposase reaction.

ATAC-seq was performed as previously described with slight modifications (*62*). Briefly, 3250 to 16,000 nuclei were directly treated with Tn5 transposase at 37°C for 30 min. After the enzymatic reaction, the DNA was purified with Zymo Column DNA Clean-5 and enriched by 13 cycles of PCR. The libraries were recovered from PCR by purification followed by size selection (180 to 750 bp) with KARA Pure Beads and were paired-end sequenced using an Illumina NovaSeq6000 Flowcell SP-PE50. Tn5 tagmentation, DNA purification, library preparation, and bioinformatic analysis were performed at the Genomics Platform of the Institut de Recherches Cliniques de Montréal. Sequencing reactions were performed at the Centre d'Expertise et de Services Génome Québec sequencing platform.

In vitro BMDM assays

BMDMs were starved for 6 hours and subsequently treated with either the TLR4 inhibitor

TAK-242 (final concentration 1 μM) or dimethyl sulfoxide (DMSO) (vehicle) for 1 hour. After pretreatment, FA-free BSA (5%), palmitate (500 μM), or SA (200 μM) was added to the medium for 24 hours. The medium was then refreshed and changed periodically for 5 days. Before collecting the cell lysates, BMDMs were treated with LPS (10 ng/ml) for 4 hours or HMGB1 (500 ng/ml) for 6 hours.

Seahorse extracellular flux analyzer assays

Real-time analysis ECAR and OCR were performed on BMDM from BSA-pretreated and SA-pretreated mice with the XFe96 Flux Analyzer (Seahorse Bioscience) as per the manufacturer's instructions. BMDMs were plated in a 96-well Seahorse plate at a seeding density of 1.2×10^5 cells per well the day before experiments. Cells were incubated in GlycoStress test assay medium (DMEM 5030, 2 mM glutamine; catalog no. 25030-081; Life Technologies), pH 7.4, 1 hour before ECAR measurements. Glucose (25 mM; catalog no. G7528-250G; Sigma-Aldrich), oligomycin (1 mM, catalog no. O4876; Sigma-Aldrich), and 2-deoxy-glucose (100 mM; catalog no. sc-202010A; Santa Cruz Biotechnology) were added sequentially. Glycolysis and glycolytic capacity were calculated as ECAR. For OCR, cells were incubated in MitoStress test assay medium containingDMEM 5030 added with 25 mM glucose (catalog no. G7528-250G; Sigma-Aldrich), 2 mM glutamine (catalog no. 25030-081; Life Technologies), and 1 mM NaPyr (catalog no. P5280-25G; Sigma-Aldrich) 1 hour before the experiment. Oligomycin (1 mM; catalog no. O4876; Sigma-Aldrich), carbonyl cyanide p-trifluoromethoxyphenylhydrazone (2 mM; catalog no. ab120081; Abcam), and 1 mM rotenone-antimycin (catalog no. A8674; Sigma-Aldrich) were added sequentially. Basal respiration and maximal respiration were calculated as OCR. Upon completion of Seahorse analysis, cells were lysed and their protein content quantified. OCR and ECAR data were normalized to protein content.

QUANTIFICATION AND STATISTICAL ANALYSIS

ATAC-seq analysis

After read quality control with FASTQC version 0.11.8, alignment was performed using Bowtie2 version 2.2.6 (*63*) on the GRCm38 mouse reference genome. Alignments were post-processed to remove PCR duplicates (Picard tool version 2.4.1) and reads mapping to mitochondrial DNA (samtools version 1.8). To represent the real Tn5 transposase-binding sites of 9 bp, the coordinates of the reads were shifted by 4 bp for the plus strand and by 5 bp for the minus strand using Deeptools 3.0.1 (*64*). The former was also used to remove ENCODE's blacklisted regions (signal artifact regions) and to convert Bam files to BEDPE format. MACS2 was used to identify significant peaks using a q-value < 0.05 (*65*). The Diffbind version 2.10.0 R package was used to generate the count matrix of Tn5 insertion site numbers for each consensus peak (peaks that were present in at least two samples). DARs were identified between conditions using the DESeq2 R package with adjusted $P < 0.05$ and |fold change|>2 (*66*). A positive DAR value means that the chromatin is more open and a negative DAR value means that the chromatin is more closed in the HFD-RD or HFD samples compared with RD-RD samples. DARs were annotated with their closest feature, and different transcription-binding motifs were identified with HOMER version 4.8.0.

Statistical analysis

Data are expressed as mean ± SEM unless otherwise indicated. Multiple Student's *t* tests, two-tailed Student's *t* test, or ANOVA *t* test were used to compare two or more than two groups, respectively. Statistical significance was considered when $P < 0.05$ and is indicated as follows in the figures: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. All experiments were repeated at least three times except for the ATAC-seq, which was performed once. The number of experimental replicates is indicated for each experiment in the figure legends.

REFERENCES AND NOTES

1. H. R. Coleman, C. C. Chan, F. L. Ferris 3rd, E. Y. Chew, *Lancet* **372**, 1835–1845 (2008).
2. A. Swaroop, E. Y. Chew, C. Bowes Rickman, G. R. Abecasis, *Annu. Rev. Genomics Hum. Genet.* **10**, 19–43 (2009).
3. W. L. Wong *et al*., *Lancet Glob. Health* **2**, e106–e116 (2014).
4. J. M. Seddon, R. Reynolds, Y. Yu, M. J. Daly, B. Rosner, *Ophthalmology* **118**, 2203–2211 (2011).
5. J. W. Crabb *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14682–14687 (2002).
6. C. A. Curcio, C. L. Millican, T. Bailey, H. S. Kruth, *Invest. Ophthalmol. Vis. Sci.* **42**, 265–274 (2001).
7. R. B. Thompson *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1565–1570 (2015).
8. D. H. Anderson, R. F. Mullins, G. S. Hageman, L. V. Johnson, *Am. J. Ophthalmol.* **134**, 411–431 (2002).
9. X. Guillonneau *et al*., *Prog. Retin. Eye Res.* **61**, 98–128 (2017).
10. G. S. Hageman *et al*., *Prog. Retin. Eye Res.* **20**, 705–732 (2001).
11. H. Xu, M. Chen, J. V. Forrester, *Prog. Retin. Eye Res.* **28**, 348–368 (2009).
12. K. Rashid, I. Akhtar-Schaefer, T. Langmann, *Front. Immunol.* **10**, 1975 (2019).
13. N. Joachim, P. Mitchell, G. Burlutsky, A. Kifley, J. J. Wang, *Ophthalmology* **122**, 2482–2489 (2015).
14. R. Klein, T. Peto, A. Bird, M. R. Vannewkirk, *Am. J. Ophthalmol.* **137**, 486–495 (2004).
15. L. S. Lim, P. Mitchell, J. M. Seddon, F. G. Holz, T. Y. Wong, *Lancet* **379**, 1728–1738 (2012).
16. V. Behnke, A. Wolf, T. Langmann, *Cell. Mol. Life Sci.* **77**, 781–788 (2020).
17. M. K. Adams *et al*., *Am. J. Epidemiol.* **173**, 1246–1255 (2011).
18. P. Haas, K. E. Kubista, W. Krugluger, J. Huber, S. Binder, *Acta Ophthalmol.* **93**, 533–538 (2015).
19. J. M. Seddon, R. Widjajahakim, B. Rosner, *Invest. Ophthalmol. Vis. Sci.* **61**, 32 (2020).
20. J. B. Dixon *et al*., *JAMA* **299**, 316–323 (2008).
21. A. Golay *et al*., *Int. J. Obes.* **9**, 181–191 (1985).
22. S. D. Long *et al*., *Diabetes Care* **17**, 372–375 (1994).
23. J. Olefsky, G. M. Reaven, J. W. Farquhar, *J. Clin. Invest.* **53**, 64–76 (1974).
24. T. Deng, C. J. Lyon, S. Bergin, M. A. Caligiuri, W. A. Hsueh, *Annu. Rev. Pathol.* **11**, 421–449 (2016).
25. A. Sene *et al*., *Cell Metab.* **17**, 549–561 (2013).
26. E. M. Andriessen *et al*., *EMBO Mol. Med.* **8**, 1366–1379 (2016).
27. V. Lambert *et al*., *Nat. Protoc.* **8**, 2197–2211 (2013).
28. O. Gavrilova *et al*., *J. Clin. Invest.* **105**, 271–278 (2000).
29. S. Klebanov, C. M. Astle, O. DeSimone, V. Ablamunits, D. E. Harrison, *J. Endocrinol.* **186**, 203–211 (2005).
30. C. N. Lumeng, J. L. Bodzin, A. R. Saltiel, *J. Clin. Invest.* **117**, 175–184 (2007).
31. Q. Wang, H. Wu, *Front. Immunol.* **9**, 2509 (2018).
32. H. Wu *et al*., *Circulation* **115**, 1029–1038 (2007).
33. B. F. Zamarron *et al*., *Diabetes* **66**, 392–406 (2017).
34. S. A. Amici *et al*., *Front. Immunol.* **9**, 1593 (2018).
35. C. Erridge, *J. Leukoc. Biol.* **87**, 989–999 (2010).
36. H. Shi *et al*., *J. Clin. Invest.* **116**, 3015–3025 (2006).
37. B. Vandanmagsar *et al*., *Nat. Med.* **17**, 179–188 (2011).
38. H. Wen *et al*., *Nat. Immunol.* **12**, 408–415 (2011).
39. J. Y. Lee, K. H. Sohn, S. H. Rhee, D. Hwang, *J. Biol. Chem.* **276**, 16683–16689 (2001).
40. K. Ganeshan, A. Chawla, *Annu. Rev. Immunol.* **32**, 609–634 (2014).
41. A. J. Freemerman *et al*., *J. Biol. Chem.* **289**, 7884–7896 (2014).
42. N. Vo, R. H. Goodman, *J. Biol. Chem.* **276**, 13505–13508 (2001).
43. W. Zhao *et al*., *J. Immunol.* **186**, 3173–3179 (2011).
44. A. J. Bannister, T. Oehler, D. Wilhelm, P. Angel, T. Kouzarides, *Oncogene* **11**, 2509–2514 (1995).
45. L. N. Tsai, T. K. Ku, N. K. Salib, D. L. Crowe, *Mol. Cell. Biol.* **28**, 4240–4250 (2008).
46. D. Lara-Astiaso *et al*., *Science* **345**, 943–949 (2014).
47. I. Goren *et al*., *Am. J. Pathol.* **175**, 132–147 (2009).
48. F. Sennlaub *et al*., *EMBO Mol. Med.* **5**, 1775–1793 (2013).
49. L. G. Fritsche *et al*., *Nat. Genet.* **48**, 134–143 (2016).
50. M. G. Netea *et al*., *Science* **352**, aaf1098 (2016).
51. A. Christ *et al*., *Cell* **172**, 162–175.e14 (2018).
52. G. Boden, *Endocrinol. Metab. Clin. North. Am.* **37**, 635–646 (2008).
53. D. Hwang, *FASEB J.* **15**, 2556–2564 (2001).
54. E. K. Anderson, A. A. Hill, A. H. Hasty, *Arterioscler. Thromb. Vasc. Biol.* **32**, 1687–1695 (2012).
55. M. G. Netea *et al*., *Nat. Rev. Immunol.* **20**, 375–388 (2020).
56. J. Thompson Legault *et al*., *Cell Rep.* **13**, 981–989 (2015).
57. V. Lamantia *et al*., *J. Nutr.* **149**, 57–67 (2019).
58. H. Zhou *et al*., *Front. Immunol.* **9**, 3082 (2019).
59. M. A. Phillips *et al*., *BMC Dermatol.* **4**, 2 (2004).
60. A. Baena *et al*., *PLOS ONE* **2**, e621 (2007).
61. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).
62. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, *Nat. Methods* **10**, 1213–1218 (2013).
63. B. Langmead, S. L. Salzberg, *Nat. Methods* **9**, 357–359 (2012).
64. F. Ramírez *et al*., *Nucleic Acids Res.* **44**, W160–W165 (2016).
65. Y. Zhang *et al*., *Genome Biol.* **9**, R137 (2008).
66. S. Anders, W. Huber, *Genome Biol.* **11**, R106 (2010).

ACKNOWLEDGMENTS

We thank all other laboratory members for helpful discussions; M. Dupuis, E. Massicotte, and M. Faquette for cell sorting and reagent preparation; the animal care technicians for mice husbandry; and O. Neyret from the Institut de Recherches Cliniques de Montréal for help with ATAC-seq. **Funding:** P.S. holds the Wolfe Professorship in Translational Research, a Canada Research Chair in Retinal Cell Biology, and is the Fonds de Recherche en Ophtalmologie de l'Université de Montréal (FROUM) Endowed Chair. Mas.H. holds the Banting Fellowship from the CIHR and Fellowship from Japan Society for the Promotion of Science (JSPS). This work was supported by operating grants to P.S from the Canadian Institutes of Health Research (foundation grant 353770), an Alcon Research Institute Senior Investigator Award, Diabetes Canada (grant DI-3-18-5444-PS), and The Heart and Stroke Foundation of Canada (grant G-16-00014658) and the BrightFocus Foundation (grant M2022015I). Additional support was provided by the FROUM and the Réseau en Recherche en Santé de la Vision. This work also benefited from infrastructure and personal support by the Canadian Foundation for Innovation (grant 20415) and the Montreal Heart Institute Foundation (C.D.R.). R.D.-M. is supported by a research scholarship from FROUM. S.C.-G. holds a Fonds de Recherche Santé du Québec (FRQS) scholarship. J.-S.J. was supported by the Canadian Institute of Health Research (CIHR grant 390615), the National Sciences and Engineering Research Council of Canada (NSERC grant 06743), and the FRQS. **Author contributions:** Mas.H., E.M.M.A.A., J.-S.J., and P.S. designed the research. Mas.H., E.M.M.A.A., Mak.H., R.D.-M., F.F., G.B., R.J., F.P., A.D., V.G., E.H., C.D., and A.M.W. performed the research. Mas.H., E.M.M.A.A., Mak.H., G.B., E.H., C.D., C.D.R., J.-S.J., and P.S. analyzed the data. Mas.H. and P.S. wrote the paper with contributions from E.M.M.A.A.,

S.C.-G., C.D.R., H.J.M., T.L., and A.M.W. **Competing interests:** P.S. and Mas.H have filed for intellectual property on the concepts presented in this study. The remaining authors declare no competing interests. **Data and materials availability:** All ATAC-seq data for the study have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus and are accessible through GEO series accession no. GSE175614. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS
science.org/doi/10.1126/science.abj8894
Figs. S1 to S12
Tables S1 to S3
MDAR Reproducibility Checklist

Submitted 9 June 2021; resubmitted 13 April 2022
Accepted 3 November 2022
10.1126/science.abj8894

EXPATRIATE SCHOLARS

Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists?

Dongbo Shi[1]*, Weichen Liu[2], Yanbo Wang[3]*

In this study, we examined China's Young Thousand Talents (YTT) program and evaluated its effectiveness in recruiting elite expatriate scientists and in nurturing the returnee scientists' productivity. We find that YTT scientists are generally of high caliber in research but, as a group, fall below the top category in pre-return productivity. We further find that YTT scientists are associated with a post-return publication gain across journal-quality tiers. However, this gain mainly takes place in last-authored publications and for high-caliber (albeit not top-caliber) recruits and can be explained by YTT scientists' access to greater funding and larger research teams. This paper has policy implications for the mobility of scientific talent, especially as early-career scientists face growing challenges in accessing research funding in the United States and European Union

Immigrants are playing an increasing role in US science and engineering (*1*–*3*), and China particularly has been the top sender of international students to the US's STEM programs (*4*, *5*). In recent years, China launched an ambitious Thousand Talents Program (TTP) to recruit elite expatriate scientists to return to China. This program has received intense attention both from the US government, as reflected in the launch of the China Initiative, and from the academic community, especially over the Federal Bureau of Investigation's arrest of Massachusetts Institute of Technology professor Gang Chen.

Despite the attention, there has been little evidence-based research on the operation, impact, and policy implications of China's talent programs. Prior research on scientific returnees has found productivity declines among those returning to lower-income home countries, but such research has focused on countries other than China (*6*). China-specific studies have suggested that returnees face difficulties reintegrating into the country's research environment (*7*), where administrative intervention and personal connections hinder scientific inquiry (*8*). This raises the question of whether China's talent programs have been effective in recruiting top-caliber scientists (*9*) and in nurturing the returnees' productivity (*10*).

Studying talent programs is important for understanding the evolving landscape of global knowledge production; it is also policy-relevant because an increasing number of governments across both high-income (e.g., Canada and Singapore) and middle- or lower-income (e.g., Brazil and India) countries are pursuing means to tap expatriates and migrant networks for domestic knowledge production and talent development. Some governments have come to believe that expats and returnees are the key to building globally competitive research institutions and dynamic, knowledge-based economies.

China's Young Thousand Talents program

We examined China's Young Thousand Talents (YTT) program, the "youth" branch of the TTP. Among the country's 200-plus talent recruitment programs, TTP is the most prominent initiative to bring leading global scientists to China. In principle, TTP is open to researchers of any nationality; but in practice, few non-Chinese have availed themselves of the program.

Established in 2010, the YTT program targets outstanding young STEM scholars and offers generous financial support to each awardee, including a one-off tax-exempt income subsidy of 500,000 yuan RMB (~$150,200 in 2010 USD purchasing power parity) and start-up grants of 1 million to 3 million yuan. This package is matched by the host institution and even local governments. All awardees are also provided with fringe benefits such as housing subsidies and are prioritized when applying for local and national grants.

To be eligible, YTT applicants should ideally (i) work in the STEM field and be 40 years of age or younger, (ii) have a PhD from a reputable overseas university and three or more years of overseas research experience, (iii) have a full-time overseas research position, (iv) be committed to full-time employment in China, and (v) be a top talent in their cohort and have the potential to become a research-field leader. However, these criteria are not rigid. YTT also welcomes freshly minted overseas PhDs and expatriate researchers with Chinese PhDs to apply if they have outstanding research records.

China's YTT program made 3576 offers between 2011 and 2017. Although designed to improve China's prospect of becoming a global STEM leader, the program's effectiveness in attracting top talents and nurturing their productivity is unclear. On one hand, a program providing substantive research support could motivate expatriate talents to return and even help grow their productivity; on the other hand, returnees may struggle to reintegrate into China's academia (*7*–*9*) and thus experience a research output slowdown. YTT scientists may also be incentivized to focus on publication quantity rather than quality, because program officials have motivations to demonstrate YTT's impact on publication counts, even at the cost of quality and originality.

Data and methods

We studied the YTT program's first four cohorts, totaling 721 awardees. Our main analyses excluded 309 individuals because they either returned for nonacademic jobs (27), received PhDs in China (196), were not of Chinese origin (34), left China within 5 years of returning (5), or lacked CV information (47). This left us with 73 scientists who rejected the YTT offers to remain overseas (hereafter, "rejectors") and 339 returnees who received PhDs abroad, accepted the offers, and spent at least 5 years conducting research in China (hereafter, "acceptors").

We implemented two sets of analyses. First, we examined YTT returnees' educational credentials and pre-return productivity. We specifically used rejector-versus-acceptor comparisons

[1]School of International and Public Affairs, Shanghai Jiao Tong University, Shanghai, China. [2]School of Public Policy and Management, Tsinghua University, Beijing, China. [3]Faculty of Business and Economics, The University of Hong Kong, Hong Kong.
***Corresponding author. Email: shidongbo@sjtu.edu.cn (D.S.); yanbo.wang@hku.hk (Y.W.)**

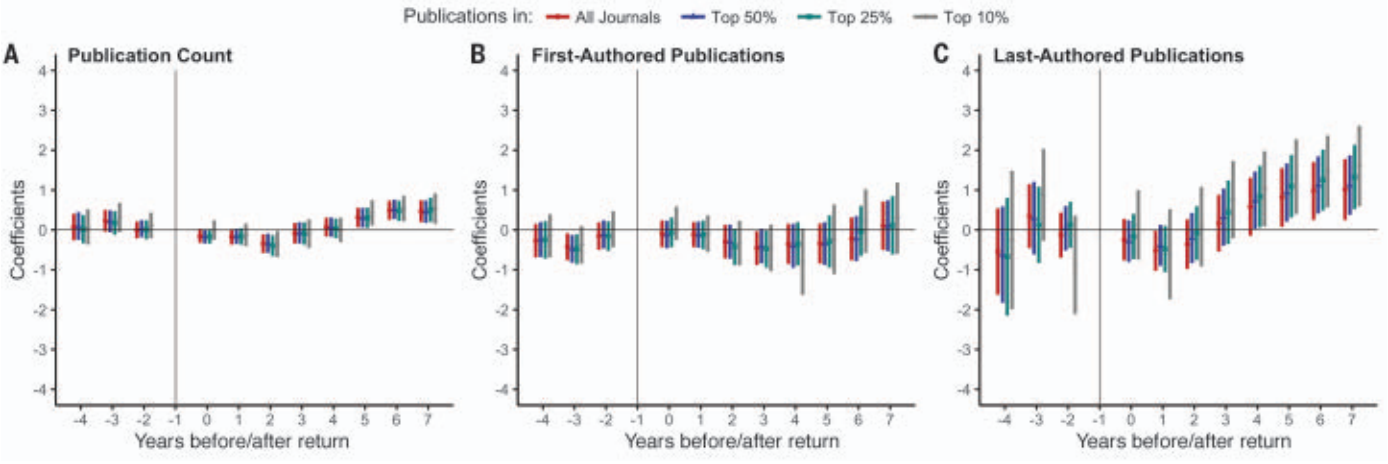


Fig. 1. Publication trajectories of YTT scientists and their overseas counterparts. The *y* axis reports coefficients estimated from Poisson regressions comparing the knowledge productivity of returnee scientists with that of their overseas counterparts for the CEM sample. The publication data are lagged by 2 years to take into consideration the necessary delay between knowledge production and in-print publication. (**A**) Annual article count without differentiating authorship position in publication. (**B** and **C**) First- and last-authored publications, respectively. The bars represent the 95% confidence intervals of the estimates. The sample includes (i) 151 returnee scientists who attended colleges in China, received their PhDs overseas, accepted the YTT offers, and spent at least 5 years of their professional careers in China and (ii) 340 overseas counterparts who had similar pre-return knowledge productivity and educational backgrounds as the YTT returnees (i.e., having attended colleges in China, received PhDs overseas, and graduated from the same doctoral institutions in the same fields around the same time period) but have stayed in overseas academia rather than returning to China.

Table 1. YTT offer receiver comparison. This table compares 339 YTT offer acceptors who have returned to China with 73 YTT offer rejectors who stayed overseas.

Covariates	Mean		Difference	*P* value
	Acceptors	Rejectors		
PhD from globally top-100 STEM program	0.525	0.551	−0.026	0.701
Research productivity before return				
Articles per year	2.390	2.932	−0.541	0.098
First-authored articles per year	1.003	1.058	−0.055	0.658
First-authored articles in top 10% of journals per year	0.523	0.403	0.119	0.040
Last-authored articles per year	0.196	0.608	−0.412	0.000
Last-authored articles in top 10% of journals per year	0.046	0.202	−0.156	0.001
Overseas faculty appointments	0.136	0.890	−0.755	0.000
Research funding per year ($1000 in 2010 USD)	4.439	30.365	−25.925	0.006

to estimate the YTT program's relative attractiveness to scientists across different levels of research caliber and career opportunity. To further contextualize YTT returnees' research capability, we benchmarked them against all early-career, research-active scientists based in the US with Chinese surnames.

Next, we evaluated the YTT program's impact on returnees' productivity using a selection-on-observables approach. We matched each returnee with comparable "stayers," that is, scientists who attended college in China and received a PhD in the same field from the same overseas university (*11*) during the same period (± 3 years) as the returnees but remained in overseas academia. To ensure YTT and stayer comparability, we further collected data on their publications and citations and used coarsened exact matching (CEM) to identify matched pairs (*12*).

We used the difference-in-differences (DID) method to estimate the YTT program's impact on the productivity of overseas-educated Chinese scientists who returned. We ran Poisson models because the outcomes are publication counts. The supplementary materials provide more details about the data and methods, and, below, we report the empirical results.

High- but not top-caliber recruits

China's YTT program has attracted high-caliber researchers, more than half of whom received PhDs from globally top-100 STEM programs (fig. S3). These recruits were also highly productive, averaging 2.39 publications annually in the pre-return period. When benchmarked against all early-career, research-active scientists based in the US (with Chinese surnames), these scientists, as a group, would rank in the top-15th (17th) percentile for productivity (table S10). The majority (73%) of the YTT recruits worked overseas as postdocs or research fellows.

The program was less successful in recruiting top-caliber scientists. Among YTT offer receivers, the rejectors were more productive (2.93 versus 2.39 publications per year; top-10th versus top-15th percentile in ranking), more likely (89% versus 14%) to have overseas faculty appointments, and associated with larger annual research grants ($30,365 versus $4439 in 2010 USD) than were the acceptors in the pre-return period (Table 1 and table S10). Furthermore, the more top-journal, first-authored publications that an offer-receiver had, the

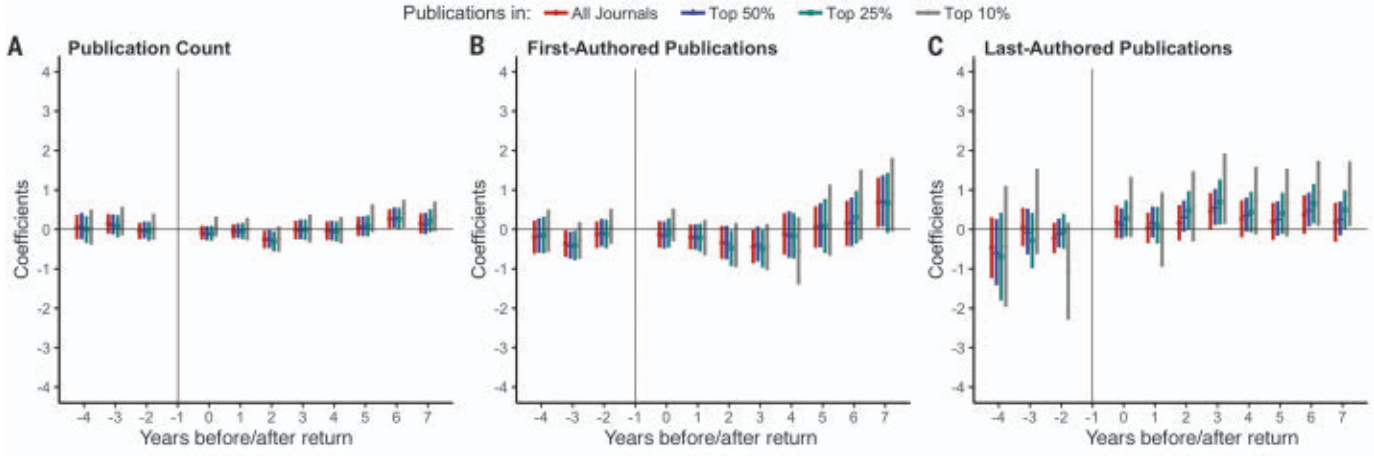


Fig. 2. Publication trajectories of YTT scientists and their overseas counterparts, controlling for team size and research funding. We obtained grant information from the Dimensions database and proxied for a scientist's team size by the annual number of unique coauthors that were listed on the scientist's last-authored publications and affiliated with their research institution. (**A**) Annual article count without differentiating authorship position in publication. (**B** and **C**) First- and last-authored publications, respectively.

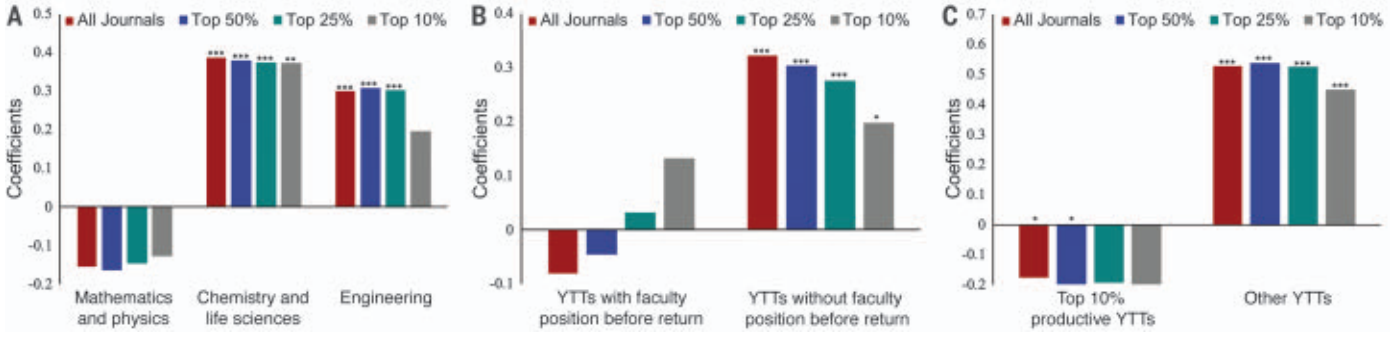

A All Journals Top 50% Top 25% Top 10%
Coefficients
Mathematics and physics
Chemistry and life sciences
Engineering
B All Journals Top 50% Top 25% Top 10%
YTTs with faculty position before return
YTTs without faculty position before return
C All Journals Top 50% Top 25% Top 10%
Top 10% productive YTTs
Other YTTs

Fig. 3. Effects of the YTT program across academic fields and scientist groups. The bars represent the difference-in-differences (DID) coefficients for the coarsened exact matching (CEM) sample across academic fields (**A**), pre-return overseas faculty appointments (**B**), and pre-return scientific productivity ranking (**C**). *P < 0.1, **P < 0.05, and ***P < 0.01.

more likely they were to have accepted the YTT offer; in contrast, there was a negative association between top-journal, last-authored publications and offer acceptance (table S11).

These results jointly show that typical YTT returnees were of high research caliber but that their pre-return productivity was ranked right below the top-10th percentile; they held no faculty positions, worked in other people's labs, and received minimal research grants. Or, to put it differently, while "the best are yet to come" (*9*), China's YTT program was attractive to young expatriates who had the capability but not the funding to run their own labs for independent research.

Returnees' productivity gain and research independence

Figure 1A shows that despite an initial drop, YTT scientists' post-return productivity was 27.4% higher than that of their CEM-matched overseas peers in terms of total publications (table S14). As matched-group scientists averaged 3.77 publications in 2016, an additional 1.03 (27.4% of 3.77) articles would have raised their ranking in Microsoft Academic Graph publication count from the 88th to the 92nd percentile. YTT scientists' performance gain continued to hold when we looked only at publications in high-impact journals, that is, journals ranked among the top 50%, 25%, and 10% in field-specific journal impact factor.

Compared with the stayers, YTT returnees had slightly fewer first-authored post-return publications (Fig. 1B); however, returnees overperformed in last-authored post-return publications by 144.3%, and this gain held across journal-quality tiers (Fig. 1C). As the STEM fields' norm is to list the principal investigator as a publication's last and corresponding author, these results suggest that YTT returnees were more likely to become independent researchers pursuing their own scientific agendas in the post-return period than were their overseas peers. In contrast, the stayers were more likely to work in others' research groups.

Numus and *manus*

We further investigated each scientist's research funding and team size. As early-career scientists in the US and EU often lack sufficient funding (*13*), YTT scientists may have benefited from the program's generous start-up grants and China's abundant supply of STEM students.

Figure 2A shows that once funding and team size were controlled for, YTT scientists barely outperformed the control-group scientists in terms of publications. This result held for both total publications and high-impact journal publications. While YTT returnees continued to publish more last-authored articles than did the stayers, the effect size became much smaller (compare Fig. 2C and Fig. 1C). For example, the incident rate ratio dropped from 2.443 to 1.371 in overall publications (compare table S14 and table S15). These statistics suggest that funding and team size play a critical role in explaining the publication gap between the returnees and the stayers.

Heterogeneity across fields and scientists

We conducted subgroup analyses across research fields and scientist profiles in pre-return employment and productivity. Figure 3A shows that YTT returnees overperformed in the fields of chemistry and life sciences, which require

large amounts of physical assets, financial resources, and human power (*14*). YTT returnees also outperformed in environmental and earth science, engineering and material science, and information science. However, we saw performance loss (although it was not statistically significant) for returnees in the fields of mathematics and physics. Figure 3, B and C, further shows that the post-return publication boost was confined to returnees who were neither overseas faculty members nor top-ranked in pre-return productivity.

Discussion and policy implications

Our empirical results show that China's YTT program has been successful in recruiting and nurturing high-caliber scientists and that YTT scientists outperform their overseas peers in post-return publication, mainly owing to their access to greater funding and larger research teams. These results show the potential of talent programs as a policy tool for countries to attract expatriate scientists and promote their productivity.

We also find that few top-caliber scientists have availed themselves of this program, suggesting room for improvement in Chinese research institutions. With the option to pursue independent research either in the US or in China, top-caliber expatriates remain unlikely to return even given the YTT offers, probably reflecting a social and cultural environment conducive to scientific inquiry in the US (*15*). The departure of Chenyang Xu—a YTT recruit and a Breakthrough Prize winner in mathematics—back to the US has specifically raised questions about whether a research environment distinguished by administrative interventions and personal connections could be conducive to nurturing top-caliber scientists (*7*–*9*).

This study's context is noteworthy. Although generously funded, the YTT program accounted for only a small portion (0.36% in 2017) of the Chinese central government's academic research and development (R&D) budget (*16*). China has been increasing its higher education expenditure (e.g., by 23.2%, compared with a 5.7% increase in US higher education expenditure, between 2018 and 2019); given YTT's relatively small budget share and the program's success in recruiting high-caliber scientists, it is highly probable that such talent programs will be sustained or even scaled up.

This study has important implications for global academic mobility, because Chinese citizens not only account for a large share of the US and EU STEM PhD graduates but also are among the most productive graduates (*5*). As China continues to invest in higher education and academic talent, we can expect more Western-trained Chinese students to return to China, although findings were mixed; whereas a National Science Foundation survey showed that 87% of Chinese STEM PhDs wanted to stay in the US (*15*), another study revealed that 70% of them would prefer to return to China if offered salaries comparable to what they could expect to receive in the US (*17*). We can also expect Chinese universities to become more attractive locations for Chinese (and international) students intending to pursue scientific research careers—students who would otherwise study in the US or EU.

If either of the previously mentioned scenarios materializes, it may disrupt the current model of university science in the US, particularly in certain academic fields. In biomedical research, for example, the field's knowledge-production function critically hinges on a large supply of postdoctoral fellows that accept minimal compensation from these temporary positions despite facing dim prospects of finding long-term tenure-track positions (*1*, *18*). The success of talent programs in countries such as China, and possibly elsewhere, would offer science-oriented international students a viable alternative to US universities and institutions. If this trend persists, the biomedical labs in the US could be facing a shrinking pool of foreign students, raising doubts about their current research model's sustainability.

Our findings also point to the need for policy adjustments to allocate more support for young scientists. It has been documented that a declining share of research grants has been going to early-career researchers in the US and EU, such that many talented young scientists cannot get a healthy start to pursue independent research (*13*, *19*). The empirical evidence from our study underscores this issue, as the relative success of the recruits of China's talent program can largely be attributed to the availability of better funding and larger research teams supporting their research. As a major driver for researchers to stay in academia is to pursue independent research (*20*), the dearth of necessary resources in the US and EU may not only expedite expatriates' return decisions but also motivate young US- and EU-born scientists to seek international research opportunities (*21*).

REFERENCES AND NOTES

1. R. Freeman, E. Weinstein, E. Marincola, J. Rosenbaum, F. Solomon, *Science* **294**, 2293–2294 (2001).
2. G. C. Black, P. E. Stephan, in *American Universities in a Global Market*, C. T. Clotfelter, Ed. (Univ. of Chicago Press, 2010), pp. 129–161.
3. J. Bound, S. Turner, P. Walsh, "Internationalization of U.S. Doctorate Education," NBER Working Paper No. 14792 (2009); http://www.nber.org/papers/w14792.
4. National Science Board, National Science Foundation, *Science and Engineering Indicators 2022: The State of U.S. Science and Engineering*, NSB-2022-1 (2022); https://ncses.nsf.gov/pubs/nsb20221.
5. P. Gaulé, M. Piacentini, *Rev. Econ. Stat.* **95**, 698–701 (2013).
6. S. Kahn, M. MacGarvie, *Rev. Econ. Stat.* **98**, 397–414 (2016).
7. C. Cao, D. F. Simon, in *Innovation and China's Global Emergence*, E. Baark, B. Hofman, J. Qian, Eds. (NUS Press, 2021), pp. 90–112.
8. Y. Xie, C. Zhang, Q. Lai, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9437–9442 (2014).
9. D. Zweig, S. Kang, H. Wang, *J. Contemp. China* **29**, 776–791 (2020).
10. Prior studies such as (*9*) have examined China's other talent programs (e.g., the TTP and the Changjiang Scholars Program) and found that the research quality of part-time participants was higher than that of full-time participants. Although insightful, these studies have examined neither the participants' research quality compared with that of nonparticipants nor these programs' impact on returnee productivity.
11. W. W. Ding, A. Ohyama, R. Agarwal, *Nat. Biotechnol.* **39**, 1019–1024 (2021).
12. S. Iacus, G. King, G. Porro, *Polit. Anal.* **20**, 1–24 (2012).
13. A. I. Leshner, *Science* **320**, 849–849 (2008).
14. Chinese scientists in such fields may also have a "regulatory advantage" over their US and EU peers.
15. R. Zwetsloot, J. Feldgoise, R. Dunham, "Trends in U.S. intention-to-stay rates of international Ph.D. graduates across nationality and STEM fields," Center for Security and Emerging Technology (CSET) Issue Brief (Georgetown University, 2020); https://cset.georgetown.edu/publication/trends-in-u-s-intention-to-stay-rates-of-international-ph-d-graduates-across-nationality-and-stem-fields/.
16. The central government budget for the 2017 (ninth cohort) YTT Program was 1.26 billion RMB, less than 0.36% of the total R&D budget (354.6 billion RMB) for universities and state research institutions (http://www.stats.gov.cn/tjsj/tjgb/rdpcgb/qgkjjftrtjgb/201908/t20190830_1694754.html).
17. R. Zeithammer, R. Kellogg, *J. Mark. Res.* **50**, 644–663 (2013).
18. J. Miller, M. Feldman, *Camb. J. Regions Econ. Soc.* **7**, 289–305 (2014).
19. Y. Xie, A. A. Killewald, *Is American Science in Decline?* (Harvard Univ. Press, 2012).
20. S. Stern, *Manage. Sci.* **50**, 835–853 (2004).
21. S. Williams, *Science* **354**, 644–647 (2016).
22. D. Shi, W. Liu, Y. Wang, Replication Data for: Has China's Young Thousand Talents Program been Successful in Recruiting and Nurturing Top Caliber Scientists?, version 1, Harvard Dataverse (2022); https://doi.org/10.7910/DVN/8SR0V9.

ACKNOWLEDGMENTS

We thank Z. Zhang, N. Liu, M. Li, J. Zhang, H. Zhou, and H. Zeng for capable research assistance. We also thank W. Ding, T. Stuart, E. Zuckerman, P. Gaulé, J. Li, W. Ng, I. Png, J. Bian, C. Marquis, Q. Wang, M. Peng, and the audiences from the National University of Singapore, the Taiwan Symposium on Innovation Economics and Entrepreneurship, the University of Hong Kong, Sun Yat-Sen University, Beijing Normal University, the Leibniz Centre for European Economic Research (ZEW), the 2022 Academy of Management Annual Conference, and the National Academies of Sciences, Engineering, and Medicine for feedback and suggestions. We are particularly grateful for the insightful input from M. Macgarvie, an early member of the research team, on the research design and empirical strategy. We deeply appreciate the insightful guidance provided by our editor and three anonymous reviewers. **Funding:** D.S. was supported by funding from the National Natural Science Foundation of China (grant #71704107), Shanghai Chen Guang Project (grant #17CG04), and Shanghai Research Center for Innovation and Policy Evaluation. Y.W. was supported by the National University of Singapore's Humanities and Social Sciences (HSS) Research Fellowship and the University of Hong Kong's Startup Grant. **Author contributions:** D.S. and Y.W. collaboratively conceived of and designed the study. Y.W. drafted the manuscript. D.S. and Y.W. revised and edited the manuscript. D.S. and W.L. collected and analyzed the data. D.S. implemented all the regressions and produced all visualizations. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Data for replicating the results in this paper are available in the Harvard Dataverse (*22*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq1218
Materials and Methods
Figs. S1 to S6
Tables S1 to S32
References (*23*–*56*)

Submitted 28 March 2022; accepted 30 November 2022
10.1126/science.abq1218

DEVELOPMENT

Immotile cilia mechanically sense the direction of fluid flow for left-right determination

Takanobu A. Katoh[1,2]*, Toshihiro Omori[3]*, Katsutoshi Mizuno[1]†, Xiaorei Sai[1], Katsura Minegishi[1]‡, Yayoi Ikawa[1], Hiromi Nishimura[1], Takeshi Itabashi[4], Eriko Kajikawa[1], Sylvain Hiver[1], Atsuko H. Iwane[4], Takuji Ishikawa[3], Yasushi Okada[5,6], Takayuki Nishizaka[2], Hiroshi Hamada[1]*

Immotile cilia at the ventral node of mouse embryos are required for sensing leftward fluid flow that breaks left-right symmetry of the body. However, the flow-sensing mechanism has long remained elusive. In this work, we show that immotile cilia at the node undergo asymmetric deformation along the dorsoventral axis in response to the flow. Application of mechanical stimuli to immotile cilia by optical tweezers induced calcium ion transients and degradation of *Dand5* messenger RNA (mRNA) in the targeted cells. The Pkd2 channel protein was preferentially localized to the dorsal side of immotile cilia, and calcium ion transients were preferentially induced by mechanical stimuli directed toward the ventral side. Our results uncover the biophysical mechanism by which immotile cilia at the node sense the direction of fluid flow.

The breaking of left-right (L-R) symmetry depends on a unidirectional fluid flow at the L-R organizer in fish, amphibians, and mammals (*1*, *2*) but not in reptiles and birds (*3*, *4*). In the mouse embryo, the leftward flow at the ventral node (the L-R organizer in this species) is generated by clockwise rotation of motile cilia on pit cells located in the central region of the node. This unidirectional flow is likely sensed by immotile (primary) cilia on crown cells located at the periphery of the node (*5*). How the embryo senses this fluid flow and why the left-side cilia preferentially respond have not been understood previously. Although mechanosensing and chemosensing have each been proposed to underlie this process (*6*), the precise mechanism has remained elusive, largely as a result of technical difficulties.

Asymmetric deformation of immotile cilia in response to the flow

We first examined how immotile cilia behave in response to the nodal flow in vivo. High-speed live fluorescence imaging revealed that nodal flow induces frequent small bending movements of both left- and right-side cilia (fig. S1 and movie S1), which suggests that continuous time-independent motion, rather than bilaterally equal periodic motion, contributes to L-R symmetry breaking. To examine whether immotile cilia undergo steady-state deformation in response to nodal flow, we compared the shape of the same cilium in the presence or absence of the flow. Immotile cilia at the node were labeled with mNeonGreen with the use of NDE, a crown cell–specific enhancer derived from the mouse *Nodal* gene, whereas the cytoplasm of crown cells was labeled with tdKatushka2 (Fig. 1A). Cilia labeled with mNeonGreen were located at the periphery of the node, and most of them were negative for Foxj1 (fig. S2A). Transmission electron microscopy (TEM) also revealed that cilia at the periphery of the node lacked outer dynein arms (fig. S2C), further confirming the immotility of crown cell cilia (*5*) (movie S1). Motile cilia at the node were immobilized by ultraviolet (UV) irradiation, which is thought to induce cleavage of dynein heavy chains (*7*), and the shape of immotile cilia was observed by high-resolution microscopy before and after such irradiation (Fig. 1, B and C, and fig. S3). Nodal flow, as revealed by particle image velocimetry (PIV) analysis, was completely lost after UV irradiation for 45 s (fig. S3, A and B, and movie S2). The flow-dependent bending angle of an immotile cilium was then estimated by ellipsoidal fitting of the shape of the same cilium before and after UV irradiation (Fig. 1, D to F, and movie S3). Examination of the bending angle of immotile cilia of embryos at the two-somite stage revealed that cilia on the left side bent toward the ventral side by 5.0° ± 9.2° (mean ± SD; n = 21), whereas those on the right side bent toward the dorsal side by 4.2° ± 7.4° (n = 18) (Fig. 1, G and H). This difference in bending angle depended on the presence of the leftward flow, given that it was lost in *iv/iv* mutant embryos (Fig. 1H), which lack nodal flow, and it was dependent on embryonic stage (Fig. 1H). The bending angle was thus significantly asymmetric along the dorsoventral (D-V) axis at the two- and three-somite stages, when the velocity of nodal flow is maximal, whereas it did not manifest asymmetry at the late headfold (LHF) and zero-somite stages, when the flow is absent or weak, respectively (*8*) (Fig. 1H). This asymmetric bending of immotile cilia along the D-V axis is consistent with the direction of the flow. Modeling the flow on the basis of in vivo observations thus suggested the presence of a ventrally directed flow at the left-posterior region of the node and a dorsally directed flow on the right side of the node (fig. S4 and movie S4). There was no significant asymmetry in the bending angle along the anterior-posterior (A-P) axis of embryos examined between the LHF and three-somite stages (fig. S3F). Notably, immotile cilia with the largest extent of ventral bending were preferentially observed at the left-posterior region of the node (Fig. 1G and fig. S3E), where the ventrally directed flow is prominent (fig. S4C) and the first molecular asymmetry appears (*9*).

Immotile cilia at the node respond to mechanical stimuli

Given that immotile cilia on the right and left sides of the node were found to bend asymmetrically along the D-V axis in response to the leftward fluid flow, we next tested whether immotile cilia at the node respond to mechanical force with the use of optical tweezers (*10*, *11*) (Fig. 1B and fig. S5). We examined mouse embryos harboring two transgenes (Fig. 2, A and B, and movie S5)—one to visualize perinodal immotile cilia and the other to monitor the response to mechanical stimuli. *Dand5* mRNA is the ultimate target of nodal flow (*12*, *13*), being degraded by the Bicc1-Ccr4 complex in response to the flow (*14*). Crown cells were labeled with the *NDE4-hsp-dsVenus-Dand5-3′-UTR* transgene, with the level of dsVenus mRNA reflecting that of *Dand5* mRNA (*14*), whereas their immotile cilia were visualized with a transgene encoding mCherry. We applied whole-cell fluorescence recovery after photobleaching (FRAP) to examine the kinetics of *Dand5* mRNA after subjecting an immotile cilium to mechanical stimuli (fig. S6A). We first tested the validity of the whole-cell FRAP system. According to a theoretical model, the rate of fluorescence recovery depends on the mRNA level (fig. S6B). In wild-type embryos with a leftward flow, the level of fluorescence on the right side of the node rapidly recovered after photobleaching, consistent with the predicted FRAP curve, whereas the level of recovery was much lower on the left side (fig. S6C). These observations confirmed that the whole-cell FRAP system was able to monitor the kinetics of *Dand5* mRNA.

[1]Laboratory for Organismal Patterning, RIKEN Center for Biosystems Dynamics Research, Kobe, Hyogo, Japan. [2]Department of Physics, Faculty of Science, Gakushuin University, Toshima-ku, Tokyo, Japan. [3]Graduate School of Biomedical Engineering, Tohoku University, Aoba Aramaki, Sendai, Miyagi, Japan. [4]RIKEN Center for Biosystems Dynamics Research, Higashi-Hiroshima, Hiroshima, Japan. [5]Laboratory for Cell Polarity Regulation, RIKEN Center for Biosystems Dynamics Research, Suita, Osaka, Japan. [6]Department of Cell Biology and Physics, Universal Biology Institute and International Research Center for Neurointelligence, The University of Tokyo, Hongo, Tokyo, Japan.
*Corresponding author. Email: takanobu.a.katoh@gmail.com (T.A.K.); omori@tohoku.ac.jp (T.O.); hiroshi.hamada@riken.jp (H.H.)
†Present address: Department of Cell Biology and Biochemistry, Division of Medicine, Faculty of Medical Sciences, University of Fukui, Eiheiji-cho, Yoshida-gun, Fukui, Japan.
‡Present address: Department of Molecular Therapy, National Institutes of Neuroscience, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan.

Mechanical stimuli were administered under a condition that mimics nodal flow to individual immotile cilia of *iv/iv* embryos (which lack nodal flow) at the early headfold (EHF) stage to the three-somite stage by positioning a polystyrene bead (with a diameter of 3.5 μm) trapped by optical tweezers into contact with the cilium and displacing it 1.75 μm toward the ventral side and then 1.75 μm toward the dorsal side at a frequency of 2 Hz (Fig. 2C and movie S6). The amplitude is within the physiological range (fig. S3G), and the maximal trapping force of ~±12 pN was sufficient to apply mechanical bending to an immotile cilium (fig. S5A). Observation of beads by three-dimensional (3D) single-particle tracking microscopy (*15*, *16*) confirmed that they moved along the D-V axis (fig. S5C). The infrared laser of the optical tweezers did not exert any unexpected effects, such as a change in Ca^{2+} oscillation pattern in crown cells or cilia (fig. S7 and materials and methods). After administration of mechanical stimuli to an immotile cilium for 1.5 hours, all crown cells were subjected twice to uniform photobleaching with a recovery period of 30 min after each bleaching. The timing and duration of the stimulation matched our previous in vivo observations (*8*). The recovery of dsVenus fluorescence in the cell with the stimulated cilium and in neighboring unstimulated ciliated cells was monitored by time-lapse 3D imaging (Fig. 2, D and E; fig. S6D; and movie S7). The fluorescence intensity at 30 min after each bleaching had reached a plateau and was compared between stimulated and neighboring cells (Fig. 2F and fig. S6E). The extent of fluorescence recovery in the stimulated cell was substantially lower than that in the unstimulated cells, with values of 69.9 ± 21.5% at 2 hours and 53.0 ± 20.9% at 2.5 hours after the onset of stimulation (geometric means ± SDs; n = 28) (Fig. 2G), which suggests that mechanical stimulation of an immotile cilium was able to induce degradation of *Dand5* mRNA. Immotile cilia on the right and left sides of the node of *iv/iv* embryos responded similarly to the mechanical stimuli (Fig. 2H). Those on the right side of the node of wild-type embryos (in the presence of the endogenous nodal flow) also showed a similar response to mechanical stimuli (Fig. 2I; fig. S6, G to I; and movie S8). Administration of mechanical stimuli to the cell body of a crown cell instead of to its cilium did not affect recovery of the fluorescence signal (Fig. 2G). Furthermore, the response to mechanical stimuli was lost in embryos lacking the cation channel Pkd2 (Fig. 2I), which suggests that mechanical stimulation of immotile cilia induces degradation of *Dand5* mRNA in a Pkd2-dependent manner. Mechanical stimulation for 45 min instead of 1.5 hours was sufficient to promote *Dand5* mRNA degradation (fig. S6J).

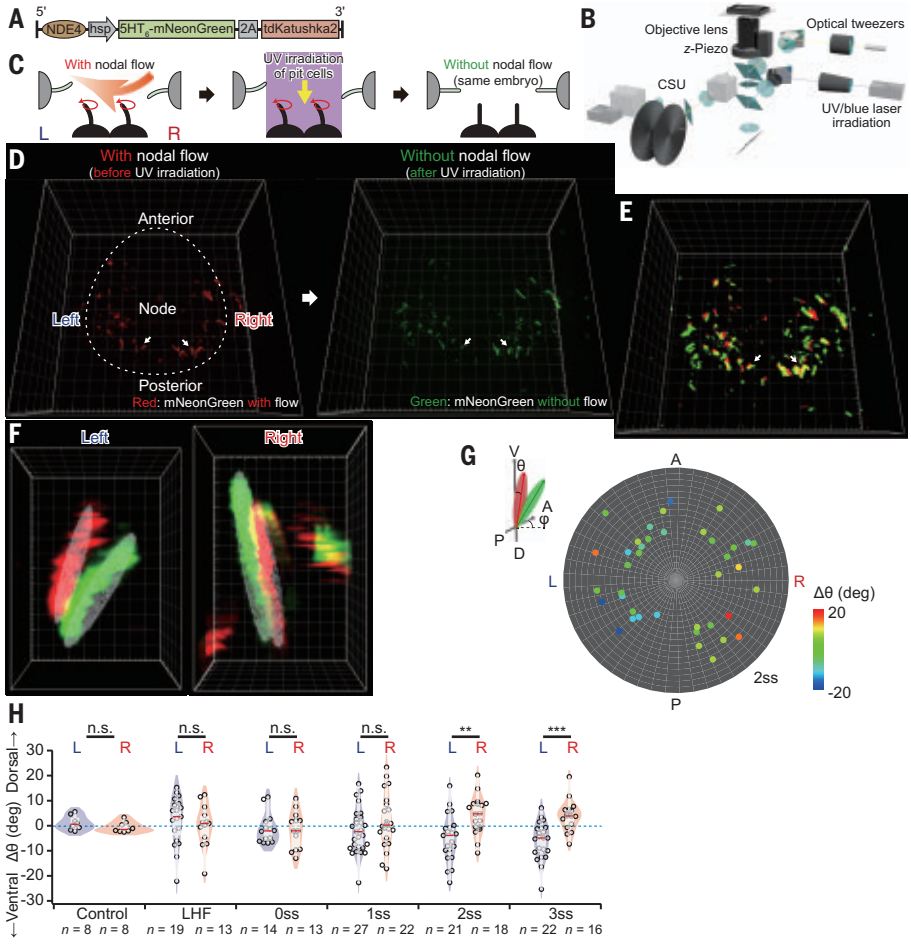


Fig. 1. Immotile cilia at the node of mouse embryos undergo asymmetric deformation along the D-V axis in response to nodal flow. (**A**) Schematic of the transgene. Immotile cilia at the node are visualized on the basis of mNeonGreen expression that is under the control of the NDE and is targeted to cilia by a 5-hydroxytryptamine receptor isoform 6 ($5HT_6$) sequence. (**B**) Schematic of the optical pathway for analysis. A UV laser and blue laser for irradiation are introduced into a microscope, which is equipped with a spinning-disk confocal unit (CSU) and optical tweezers. (**C**) Schematic of the experiment. Live fluorescence images of immotile cilia at the node were first obtained in the presence of nodal flow. The central region of the node encompassing pit cells was then subjected to UV irradiation to abolish nodal flow, and fluorescence images of immotile cilia in the absence of the flow were obtained from the same embryo. (**D**) High-resolution 3D images obtained by deconvolution processing of immotile cilia (fig. S3D) in the presence (left) or absence (right) of nodal flow. Cilia shown in red or green correspond to those in the presence or absence of the flow, respectively. Grid size, 10 μm. (**E**) Detection of the edge of each cilium after alignment. Immotile cilia of the same embryo are shown in the presence (red) and absence (green) of the flow. Grid size, 10 μm. (**F**) Individual immotile cilia on the left or right side of the node in the presence (red) or absence (green) of the flow. The zenith and azimuth angles were determined by ellipsoidal fitting (gray mesh) of the edge of each cilium. Grid size, 1 μm. (**G**) Distribution of immotile cilia at the node with various values of Δθ (change in the zenith angle in response to the flow). Data were obtained from wild-type embryos (n = 39 cilia from 7 embryos) at the two-somite (2ss) stage (fig. S3E). A, anterior; P, posterior; R, right; L, left; V, ventral; D, dorsal. (**H**) Δθ values [degrees (deg)] for immotile cilia on the left and right sides of the node were determined at various developmental stages. Data for *iv/iv* embryos at the LHF stage to the 2ss stage are shown as a control. Red bars indicate median values. $^{**}P < 0.01$; $^{***}P < 0.001$; n.s., not significant (Mann-Whitney U test).

The response to mechanical stimuli was further confirmed with another readout: expression of the transgene *ANE-LacZ*, which allows monitoring of Nodal activity in perinodal cells (*17*). This transgene manifests left-sided expression at the node in the presence of nodal flow but shows L-R randomized expression in its absence (*9*) (fig. S6F). Mechanical stimulation along the D-V axis of an immotile cilium on the right side of an *iv/iv* embryo resulted in a significant increase in the R/(R + L) ratio of *ANE-LacZ* expression to 0.64 ± 0.19 (mean ± SD; n = 12) compared with a value of 0.47 ± 0.22 (n = 24) for control embryos (Fig. 2J). These

results suggested that mechanical stimulation of a single immotile cilium not only induced degradation of *Dand5* mRNA in the targeted cell but also established molecular asymmetry in all crown cells at the node. A feedback mechanism involving Wnt and Dand5 signals (*12*) may be responsible for the expansion of asymmetric Nodal activity among crown cells.

Perinodal cells of mouse embryos manifest both cytoplasmic and intraciliary Ca^{2+} transients in response to nodal flow (*18*–*20*). We therefore examined whether mechanical stimuli administered to immotile cilia of *iv/iv* embryos might induce such transients. Cytoplasmic and intraciliary Ca^{2+} transients were observed with cytoplasm-targeted and cilium-targeted (*18*) forms of the fluorescent Ca^{2+} indicator GCaMP6, respectively (Fig. 3, A and B). Spontaneous Ca^{2+} transients were detected in both the cytoplasm and immotile cilia (Fig. 3C), with such transients having been shown to be independent of nodal flow and the Pkd2 channel (*18*, *19*). However, the frequency of Ca^{2+} transients increased significantly from 0.83 ± 0.71 to 1.39 ± 1.65 spikes per minute in the cytoplasm (means ± SDs; n = 42 cells) and from 0.32 ± 0.57 to 0.59 ± 1.02 spikes per minute in cilia (n = 24) in response to mechanical stimuli (Fig. 3, C and D, and movie S9). Such increases in the frequency of cytoplasmic and ciliary Ca^{2+} transients were not observed in embryos lacking the Pkd2 channel (Fig. 3E).

Immotile cilia sense bending direction in a manner dependent on polarized localization of Pkd2

We next investigated whether an immotile cilium might respond differentially to forced bending toward the dorsal or ventral sides, possibly as a result of a structure or molecule within the cilium that can sense the direction of bending. The polarized distribution of such a structure or molecule relative to the midline of an embryo would allow a differential response to the direction of nodal flow (fig. S8A). We first searched for such a structure at or near the base of immotile cilia by focused ion beam–scanning electron microscopy (FIB-SEM), which would be expected to reveal an

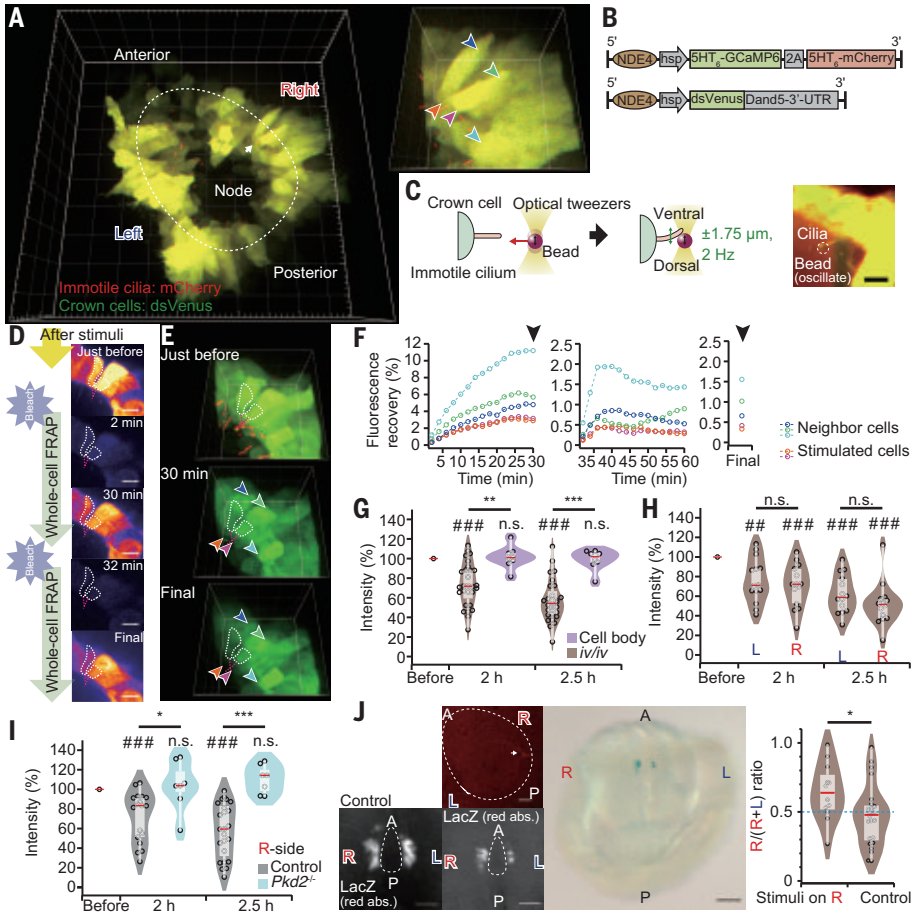


Fig. 2. Mechanical stimuli administered to immotile cilia by optical tweezers trigger *Dand5* mRNA degradation and increase Nodal activity. (**A**) A 3D image of the node of an *iv/iv* mouse embryo at the two-somite stage harboring the two transgenes shown in (B). mCherry (red) marks immotile cilia, whereas *Dand5* mRNA degradation in crown cells can be monitored by measurement of dsVenus fluorescence (green). The white arrow indicates cilia to which mechanical stimuli were applied. Grid size, 20 μm. The cells to which mechanical stimuli were applied (orange and purple arrowheads) and surrounding unstimulated cells (blue, green, and cyan arrowheads) are also shown at higher magnification to the right of the main image. Grid size, 10 μm. (**B**) Schematic of the two transgenes adopted for these experiments. UTR, untranslated region. (**C**) Experimental scheme. A polystyrene bead is trapped, placed into contact with an immotile cilium, and forced to oscillate along the D-V axis for 1.5 hours with the use of optical tweezers. The image on the right shows an oscillating bead (white dotted line) making contact with cilia (movie S6). (D to H) Analysis of *iv/iv* embryos at the EHF to three-somite stages. (**D**) *Dand5* mRNA degradation was monitored by whole-cell FRAP (fig. S6, A to C). The entire area of the targeted cells was bleached twice with a 30-min interval between sessions, and fluorescence recovery was monitored. White dotted lines indicate the stimulated cells. 2D sections obtained from 3D images are shown. Scale bars, 10 μm. (**E**) 3D images obtained during FRAP. Red, purple, blue, green, and cyan arrowheads represent the same cells shown in (A). Grid size, 10 μm. (**F**) Time course of fluorescence recovery during the first and second FRAP periods. A 3D image was obtained with a longer exposure time at the end of the second FRAP session (fig. S6E). Normalized intensity was calculated with the use of the values indicated by the closed arrowheads and is shown in (G). (**G**) Normalized fluorescence intensity of dsVenus is shown for before as well as 2 and 2.5 hours after stimulation (brown; n = 28 embryos). Data are also shown for cells whose cell body (instead of the cilium) was stimulated (purple; n = 8 embryos). Red bars indicate median values. (**H**) Normalized fluorescence intensity of dsVenus before and after stimulation for immotile cilia on the left and right sides (n = 14 embryos each for left-side cilia and right-side cilia). (**I**) Normalized fluorescence intensity of dsVenus for similar FRAP experiments performed with $Pkd2^{-/-}$ and control (wild-type, *iv/+*, or $Pkd2^{+/-}$) embryos (n = 7 for $Pkd2^{-/-}$ and 22 for control embryos). The experiments were performed only with immotile cilia on the right side to avoid the effect of nodal flow. (**J**) A single cilium on the right side of an *iv/iv* embryo harboring the *ANE-LacZ* transgene was subjected to mechanical stimulation for 1.5 hours (arrow in the mCherry fluorescence image shown in the upper left), cultured for ~7 hours, and then subjected to X-gal staining to detect Nodal activity (large middle panel) (fig. S6F). The level of staining was quantified as red absorbance (lower left), and the ratio of the staining level on the right side to that on the right plus left sides of the node [R/(R + L)] was determined (right) for embryos subjected to mechanical stimulation or nonstimulated (control) embryos. Scale bars, 20 μm (upper left), 50 μm (lower left), and 100 μm (middle). ##P < 0.01; ###P < 0.001 [Wilcoxon signed-rank test for comparisons of before with 2 and 2.5 hours in (G) to (I)]. *P < 0.05; **P < 0.01; ***P < 0.001 [Mann-Whitney U test in (G) to (J)].

anisotropic distribution relative to the midline, but we were not successful (fig. S8B). Consistent with this result, measurement of the flexural rigidity of immotile cilia with optical tweezers revealed no apparent difference between dorsal and ventral bending (fig. S9).

Alternatively, a mechanosensitive channel may be preferentially localized to one side (dorsal or ventral) of an immotile cilium. The most likely candidate for such a channel would be Pkd2, given that the ciliary localization of this protein is essential for the breaking of L-R symmetry (*5*, *21*). We examined the precise localization of Pkd2 within immotile cilia by 3D stimulated emission depletion (STED) microscopy of wild-type embryos harboring an *NDE2-hsp-Pkd2-Venus* transgene, which is able to rescue the defects of Pkd2-deficient mouse embryos (*5*). The super-resolution images revealed a nonuniform distribution of the Pkd2::Venus protein on each immotile cilium. The Pkd2::Venus protein thus accumulated to form clusters on the surfaces of immotile cilia, with the clusters being preferentially localized to the dorsal side (the side facing the midline of the embryo) of those on both the left and right sides of the embryo (Fig. 4A and movie S10). Analysis of the angular distribution of Pkd2::Venus on the transverse plane of the axoneme revealed that the D/(D + V) ratio of Pkd2 signal intensity was significantly biased toward the dorsal side (0.54 ± 0.12, mean ± SD; n = 50) (Fig. 4B). Preferential localization of Pkd2 on the dorsal side of immotile cilia was confirmed by confocal microscopy with the Airyscan detector. Analysis of the distance along the z axis between the center of localization of Pkd2 and that of the axoneme by Gaussian fitting revealed that the Pkd2 region was displaced toward the dorsal side by 142 ± 92 nm in cilia on the left side (mean ± SD; n = 53) and by 192 ± 203 nm in those on the right side (n = 54) (Fig. 4C; fig. S10, A to C; and movie S11), with these distances being compatible with a value of 100 nm for the radius of an axoneme. Furthermore, images of longitudinal sections of immotile cilia also confirmed the dorsal localization of Pkd2 (fig. S10E). By contrast, there was no significant enrichment of Pkd2 along the proximodistal axis of a cilium (fig. S10D). Endogenous Pkd2 also showed a similar preferential distribution to the dorsal side of immotile cilia as examined with the use of antibodies to Pkd2 (fig. S11).

Enrichment of Pkd2 at the dorsal side of a cilium could explain how immotile cilia sense the direction of nodal flow (Fig. 4D). Imposition of mechanical stimuli in a single direction and measurement of cytoplasmic Ca^{2+} transients revealed that immotile cilia showed a significantly greater response to stimuli directed toward the ventral side than to those directed toward the dorsal side (Fig. 4E, fig. S12, and movie S12). Dorsal bending generated a greater response than ventral bending regardless of the order of bending direction (fig. S12, C and D). Examination of the Pkd2 expression pattern by generation of *Pkd2*^{mNG} mice, in which the amino acid sequence for mNeonGreen was knocked in at the COOH-terminus of Pkd2, revealed that the Pkd2::mNeonGreen protein was present mostly in immotile cilia of crown cells (fig. S2B).

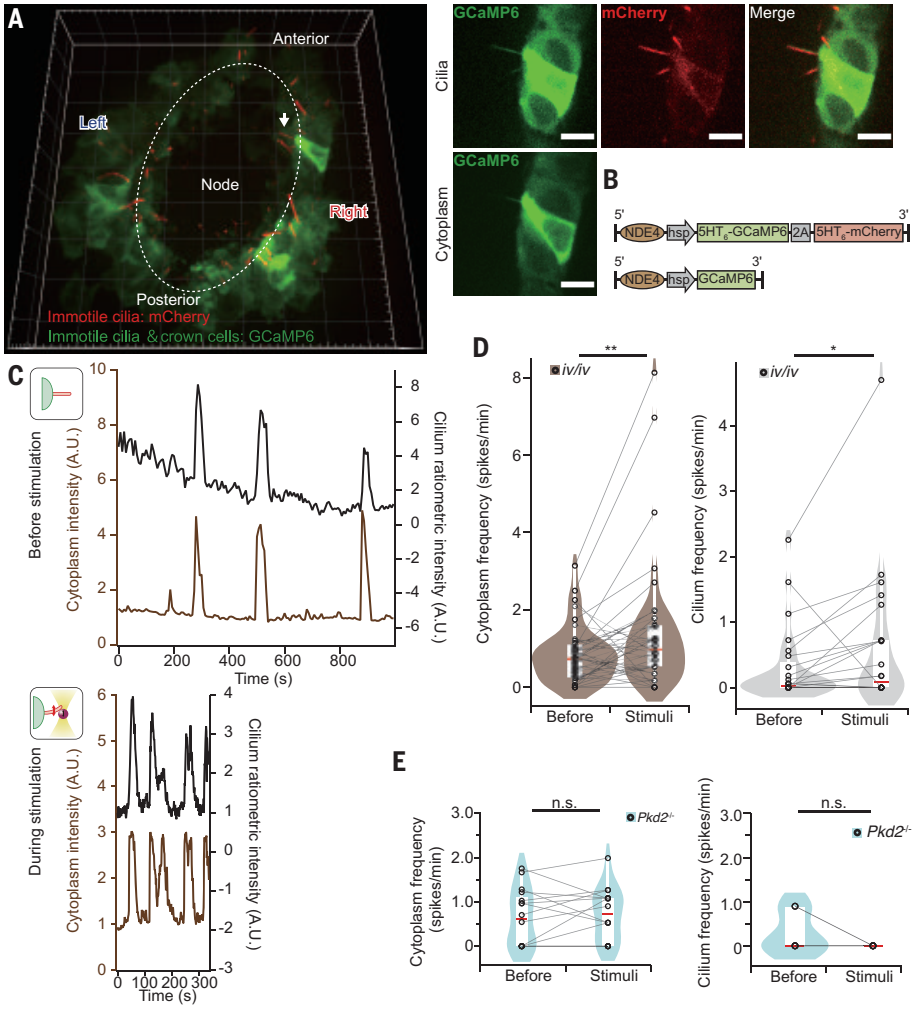


Fig. 3. Mechanical stimulation of the immotile cilium of crown cells alters the dynamics of Ca^{2+} signaling in both the cilium and cytoplasm. (**A**) A 3D image of the node of an *iv/iv* embryo at the two-somite stage harboring the two transgenes in (B) is shown on the left. The white arrow indicates a cilium to which mechanical stimuli were applied. Grid size, 10 μm. Both GCaMP6 and mCherry are expressed in immotile cilia for ratiometric Ca^{2+} imaging, with sections containing the cilium being averaged (upper right). GCaMP6 is also expressed in the cytoplasm for cytoplasmic Ca^{2+} imaging, with sections containing the cell body being averaged (lower right). Scale bars, 10 μm. (**B**) Schematic of the two transgenes used for intracellular Ca^{2+} measurement. (**C**) Time course of Ca^{2+} signal intensity before (top) and during (bottom) stimulation of the immotile cilium of a crown cell in an *iv/iv* embryo. Brown traces indicate cytoplasmic Ca^{2+} (GCaMP6 F/F_0 ratiometric values), whereas black traces indicate intraciliary Ca^{2+} (GCaMP6/mCherry F/F_0 ratiometric values). Calcium dynamics in the cilium and cytoplasm were monitored for ~15 min before (upper) and then for ~5 min after (lower) the onset of mechanical stimulation of the cilium (movie S9). A.U., arbitrary units. (**D**) The mean frequency of Ca^{2+} transients in the cytoplasm (left) and cilium (right) was measured as in (C) (n = 42 cells from 28 embryos for cytoplasm and 24 cilia from 17 embryos for cilia). (**E**) Mean frequency of Ca^{2+} transients in the cytoplasm and cilium of *Pkd2*$^{-/-}$ embryos (n = 16 cells from 16 embryos for cytoplasm and 6 cilia from 6 embryos for cilia). *P < 0.05; **P < 0.01 [Wilcoxon signed-rank test in (D) and (E)].

Discussion

Our results collectively indicate that immotile cilia at the node respond to mechanical force generated by fluid flow. This notion contradicts the previous claim that primary cilia do not function as Ca^{2+}-dependent mechanosensors (*22*) but is supported by similar findings with zebrafish embryos in an accompanying paper (*23*). Given that the relative extent of viscous

and bending force is approximately given by length^4/stiffness$^{0.25}$ (*24*), the slender shape of a cilium is suited to sensing a weak flow and transducing the flow signal into strong locoregional strain. In the presence of leftward flow, the bending of an immotile cilium on the left side of the node toward the ventral side imposes a strain of 0.014 ± 0.013 (mean ± SD; n = 8) to the dorsal side of the cilium (Fig. 4, F and G). The resulting membrane tension, according to a previously described model (*25*), would be 1.6 ± 1.6 mN/m (mean ± SD; n = 8) (fig. S13), which may be sufficient to activate dorsally localized Pkd2 and trigger the Ca^{2+} response. By contrast, on the right side of the node, strain at the dorsal side of an immotile cilium is as small as 0.000 ± 0.001 (n = 8) and would not support a response (Fig. 4, F and G).

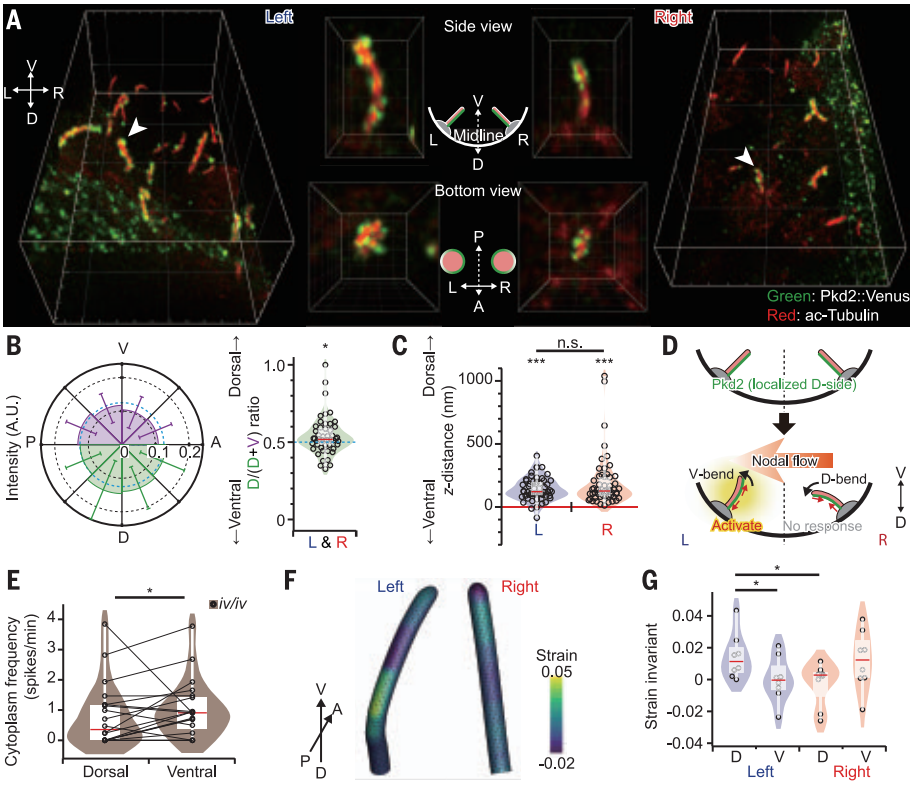


Fig. 4. Immotile cilia sense bending direction in a manner dependent on polarized localization of Pkd2. (**A**) A wild-type mouse embryo harboring an *NDE2-hsp-Pkd2-Venus* transgene was subjected to immunofluorescence analysis for detection of the Pkd2::Venus fusion protein and acetylated (ac)–tubulin at the node with a 3D-STED microscope (left and right images). Magnified views of D-V sections and bottom views of cilia (indicated by white arrowheads in left and right images) shown in the middle suggest a preferential localization of Pkd2::Venus at the dorsal side of cilia on both the left and right sides of the node. Grid size, 5 μm (main panels), 1 μm (side views), and 0.5 μm (bottom views). (**B**) The angular distribution of green fluorescence intensity in transverse planes of each cilium imaged as in (A) was analyzed (left). The ratio of Pkd2::Venus signal intensity on the dorsal side to that on the dorsal plus ventral sides [D/(D + V)] (right) was significantly biased toward the dorsal side (n = 50 cilia from 4 embryos). *P < 0.05 (one-sample t test). (**C**) Distance along the z axis between the centers of red and green fluorescence intensity in longitudinal optical sections of each cilium obtained by an Airyscan microscope (fig. S10B) was measured by Gaussian fitting after precise correction for chromatic aberration. The intensity center for Pkd2::Venus was significantly polarized toward the dorsal side of cilia on both the left and right sides (n = 53 cilia for the left side and 54 cilia for the right side from 13 embryos). ***P < 0.001 (one-sample t test). Wilcoxon signed-rank test was used for comparison of left and right sides. (**D**) Model that would explain why immotile cilia on the left side, but not those on the right side, respond to the leftward fluid flow. (**E**) Frequency of cytoplasmic Ca^{2+} transients in individual immotile cilia subjected to both dorsal and ventral bending (n = 18 cilia from 18 *iv/iv* embryos) (fig. S12, A and B). Whereas the frequency was 0.81 ± 1.03 spikes per minute (mean ± SD) for dorsal bending, it was significantly increased to 1.02 ± 0.93 spikes per minute for ventral bending. *P < 0.05 (Wilcoxon signed-rank test). (**F**) Estimated strain at the membrane of the immotile cilia on the left and right sides of the node shown in Fig. 1F. The membrane is modeled as a 2D hyperelastic material, and the contour color indicates the second strain invariant. (**G**) Comparison of strain applied to the dorsal and ventral sides of cilia on the left or right sides of the node. The mean value of the second strain invariant was used as the basis for strain measurement. Definition of the dorsal and ventral regions is described in fig. S13B. Strains on the dorsal and ventral sides of a left-side cilium are estimated as 0.014 ± 0.013 and 0.000 ± 0.013 (means ± SDs; n = 8), respectively, whereas the corresponding values for a right-side cilium are 0.000 ± 0.001 and 0.012 ± 0.017 (n = 8), respectively. *P < 0.05 (Student's paired t test).

Our findings thus suggest how immotile cilia sense the direction of nodal flow: Directional information of the flow is geometrically converted to locoregional strain, which is integrated over the polarized area of Pkd2 localization and allows activation only of cilia on the left side, thereby giving rise to robust L-R determination. Given that other proteins, such as the channel protein Hv1 (*26*) and the structural protein LRRCC1 (*27*), show asymmetric localization within cilia and the centriole, respectively, additional molecules may be localized asymmetrically in immotile cilia at the node and render the mechanism responsible for the breaking of L-R symmetry more robust.

Several questions remain, including how Pkd2 becomes preferentially localized to one side of an immotile cilium. Crown cells on the left and right sides of the node may be polarized relative to the midline, given that the organization of centrioles in crown cells on both sides was found to be polarized along the mediolateral axis (fig. S14). An unknown signal derived from the midline of the embryo may polarize Pkd2 localization. Bone morphogenetic protein (BMP) antagonists expressed at the node are candidates for such a signal. However, treatment of embryos with exogenous BMP did not affect the localization of Pkd2 at the dorsal side of immotile cilia (fig. S10F). Characterization of the mechanism responsible for the polarized localization of Pkd2 therefore awaits further study.

REFERENCES AND NOTES

1. M. Blum, K. Feistel, T. Thumberger, A. Schweickert, *Development* **141**, 1603–1613 (2014).
2. H. Shiratori, H. Hamada, *Development* **133**, 2095–2104 (2006).
3. J. Gros, K. Feistel, C. Viebahn, M. Blum, C. J. Tabin, *Science* **324**, 941–944 (2009).
4. E. Kajikawa *et al.*, *Nat. Ecol. Evol.* **4**, 261–269 (2020).
5. S. Yoshiba *et al.*, *Science* **338**, 226–231 (2012).
6. J. H. E. Cartwright, O. Piro, I. Tuval, *Phil. Trans. R. Soc. B* **375**, 20190566 (2020).
7. B. H. Gibbons, I. R. Gibbons, *J. Biol. Chem.* **262**, 8354–8359 (1987).
8. K. Shinohara *et al.*, *Nat. Commun.* **3**, 622 (2012).
9. A. Kawasumi *et al.*, *Dev. Biol.* **353**, 321–330 (2011).
10. A. Ashkin, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4853–4860 (1997).
11. T. A. Katoh *et al.*, *Sci. Rep.* **8**, 15562 (2018).
12. T. Nakamura *et al.*, *Nat. Commun.* **3**, 1322 (2012).
13. A. Schweickert *et al.*, *Curr. Biol.* **20**, 738–743 (2010).
14. K. Minegishi *et al.*, *Nat. Commun.* **12**, 4071 (2021).
15. H. P. Kao, A. S. Verkman, *Biophys. J.* **67**, 1291–1300 (1994).
16. B. Huang, W. Wang, M. Bates, X. Zhuang, *Science* **319**, 810–813 (2008).
17. K. Yashiro *et al.*, *Genes Cells* **5**, 343–357 (2000).
18. K. Mizuno *et al.*, *Sci. Adv.* **6**, eaba1195 (2020).
19. D. Takao *et al.*, *Dev. Biol.* **376**, 23–30 (2013).
20. J. McGrath, S. Somlo, S. Makova, X. Tian, M. Brueckner, *Cell* **114**, 61–73 (2003).

21. S. Field *et al.*, *Development* **138**, 1131–1142 (2011).
22. M. Delling *et al.*, *Nature* **531**, 656–660 (2016).
23. L. Djenoune *et al.*, *Science* **379**, 71–78 (2023).
24. M. C. Lagomarsino, F. Capuani, C. P. Lowe, *J. Theor. Biol.* **224**, 215–224 (2003).
25. R. Skalak, A. Tozeren, R. P. Zarda, S. Chien, *Biophys. J.* **13**, 245–264 (1973).
26. M. R. Miller *et al.*, *Cell Rep.* **24**, 2606–2613 (2018).
27. N. Gaudin *et al.*, *eLife* **11**, e72382 (2022).

ACKNOWLEDGMENTS

We thank Y. Kiyosue for support with microscopy systems; K. Kawaguchi and members of his laboratory as well as S. Nonaka for discussions; D. Takao for support with STED imaging; Tokai Electron Microscopy, Inc., for TEM imaging; the Laboratory for Ultrastructural Research (RIKEN BDR) for technical support; the Laboratory for Animal Resource and Genetic Engineering (RIKEN, BDR) for generating the *Pkd2*^{mNG} knock-in mouse; K. Takaoka, T. Ide, H. M. Takase, and K. Shiozawa for technical advice; and T. Lange for technical assistance. **Funding:** This study was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan (no. 17H01435) and from Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST) (no. JPMJCR13W5) to H.H.; by a Grant-in-Aid (no. 21K15096) from the Japan Society for the Promotion of Science (JSPS) and by the RIKEN Special Postdoctoral Researcher Program to T.A.K.; by a grant from Precursory Research for Embryonic Science and Technology (PRESTO) of JST (no. JPMJPR2142) to T.O.; by grants from JSPS (nos. 21H04999 and 21H05308) to T.I.; and by RIKEN Cluster for Science, Technology, and Innovation Hub (RCSTI) to A.H.I. 3D-STED microscopy was supported by grants from JST (nos. JPMJMS2025-15, JPMJCR20E2, JPMJCR15G2, and JPMJCR1852) and from JSPS (nos. 19H05794 and 16H06280) to Y.O. **Author contributions:** T.A.K. and T.N. designed experiments with optical tweezers. T.A.K. performed biophysical experiments with mouse embryos and analyzed the data. T.O. and T.Is. are responsible for theoretical analysis of immotile cilia. Y.I. and H.N. generated transgenic mice. S.H. genotyped transgenic mice. K.Miz. helped with analysis of Ca^{2+} transients and highly inclined and laminated optical sheet (HILO) imaging. K.Min. helped with analysis of *Dand5* mRNA degradation. E.K. examined ANE-LacZ activity. X.S. performed immunostaining. T.It. and A.H.I. performed FIB-SEM analysis of immotile cilia. Y.O. assisted with STED analysis. T.A.K., T.O., and H.H. conceived the project and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq8148
Materials and Methods
Figs. S1 to S14
References (*28*–*46*)
MDAR Reproducibility Checklist
Movies S1 to S12

Submitted 2 May 2022; resubmitted 1 November 2022
Accepted 9 December 2022
10.1126/science.abq8148

DEVELOPMENT

Cilia function as calcium-mediated mechanosensors that instruct left-right asymmetry

Lydia Djenoune[1]†, Mohammed Mahamdeh[1]†, Thai V. Truong[2], Christopher T. Nguyen[1,3,4], Scott E. Fraser[2], Martina Brueckner[5], Jonathon Howard[6], Shiaulou Yuan[1]*

The breaking of bilateral symmetry in most vertebrates is critically dependent upon the motile cilia of the embryonic left-right organizer (LRO), which generate a directional fluid flow; however, it remains unclear how this flow is sensed. Here, we demonstrated that immotile LRO cilia are mechanosensors for shear force using a methodological pipeline that combines optical tweezers, light sheet microscopy, and deep learning to permit in vivo analyses in zebrafish. Mechanical manipulation of immotile LRO cilia activated intraciliary calcium transients that required the cation channel Polycystin-2. Furthermore, mechanical force applied to LRO cilia was sufficient to rescue and reverse cardiac situs in zebrafish that lack motile cilia. Thus, LRO cilia are mechanosensitive cellular levers that convert biomechanical forces into calcium signals to instruct left-right asymmetry.

In human, rodent, amphibian, and fish embryos, left-right (LR) asymmetry is determined at the left-right organizer (LRO) by cilia that produce and transduce directional flow of extracellular fluid into asymmetric Nodal signaling and organ laterality (*1*–*13*). However, the mechanism by which cilia sense LRO flow is unknown (*14*–*17*), resulting in two leading hypotheses: Cilia are chemosensors of morphogens carried by flow (*2*, *18*) or mechanosensors of force exerted by flow (*1*, *5*, *19*, *20*). Previous work connects asymmetric calcium signals with LRO flow and LR development (*1*, *5*, *20*, *21*). However, it has been impossible to determine whether ciliary mechanosensation or chemosensation mediates these asymmetric calcium signals due to inadequate techniques for delivery of mechanical force or chemical cues specifically to LRO cilia in vivo (*14*, *15*).

We developed and deployed an in vivo method to apply precise, localized mechanical forces onto LRO cilia in zebrafish. Our optical toolbox couples optical tweezers, light sheet microscopy, and deep learning analyses to deflect cilia and measure intraciliary calcium signaling in the zebrafish LRO. Combining these tools with longitudinal assays, we revealed that the cilium is a calcium-mediated mechanosensor that is necessary, sufficient, and instructive for LR development.

Cilia can be optically deflected

Prior studies of cilia mechanosensation applied laminar fluid flow across both the cilium and apical plasma membrane, limiting identification of a cilia-specific mechanical signal (*22*–*25*). Further, laminar flow approaches may transport chemical cues in addition to mechanical stimulation. To deflect cilia in a specific and controllable manner in vivo, we devised a strategy that employs optical tweezers to apply local mechanical forces onto cilia. The ability of optical tweezers to use light to exert and measure mechanical forces onto structures ranging in size from microns to nanometers makes them ideally suited for single cells, proteins, and molecules (*26*–*28*). We employed optical tweezers to bend cilia by directly trapping them and applying force without the need for tethered beads or microscope stage movements (which may confound ciliary signaling). Our integration of optical tweezers and fluorescence microscopy in a custom-built instrument permitted simultaneous recording with cilia-targeted fluorescent calcium reporters. This enabled fully programmable, precise spatiotemporal control of ciliary bending in both zebrafish embryos and cultured cells, and exact recapitulation of in vivo physiological conditions.

We validated our cilia deflection approach on LLC-PK1 porcine renal epithelial cells, as they have long immotile cilia and are amenable to transfection (*20*). We trapped the distal tip of the cilium by focusing the laser onto cilia of wild-type (WT) LLC-PK1 cells, and mechanically bent the cilium by steering the trapping laser with piezoelectric-actuated mirrors (fig. S1, A and B, and movie S1). All deflection parameters (frequency, displacement, angularity, direction, and duration) were precisely and remotely controlled (fig. S1, C and F). To recapitulate physiological conditions in which primary cilia are subjected to and bent by pulsatile shear fluid flow, such as in the kidney ducts and vasculature (*29*, *30*), our optical tweezers were programmed to apply oscillatory deflection patterns to cilia at defined bending frequencies. Transfection

[1]Cardiovascular Research Center, Cardiology Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02129, USA. [2]Translational Imaging Center, University of Southern California, Los Angeles, CA 90089, USA. [3]Cardiovascular Innovation Research Center, Heart, Vascular, and Thoracic Institute, Cleveland Clinic, Cleveland, OH 44195, USA. [4]Division of Health Science Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [5]Departments of Pediatrics and Genetics, Yale University School of Medicine, New Haven, CT 06520, USA. [6]Department of Molecular Biochemistry and Biophysics, Yale University School of Medicine, New Haven, CT 06520, USA.
***Corresponding author. Email: shiaulou.yuan@mgh.harvard.edu**
†These authors contributed equally to this work.

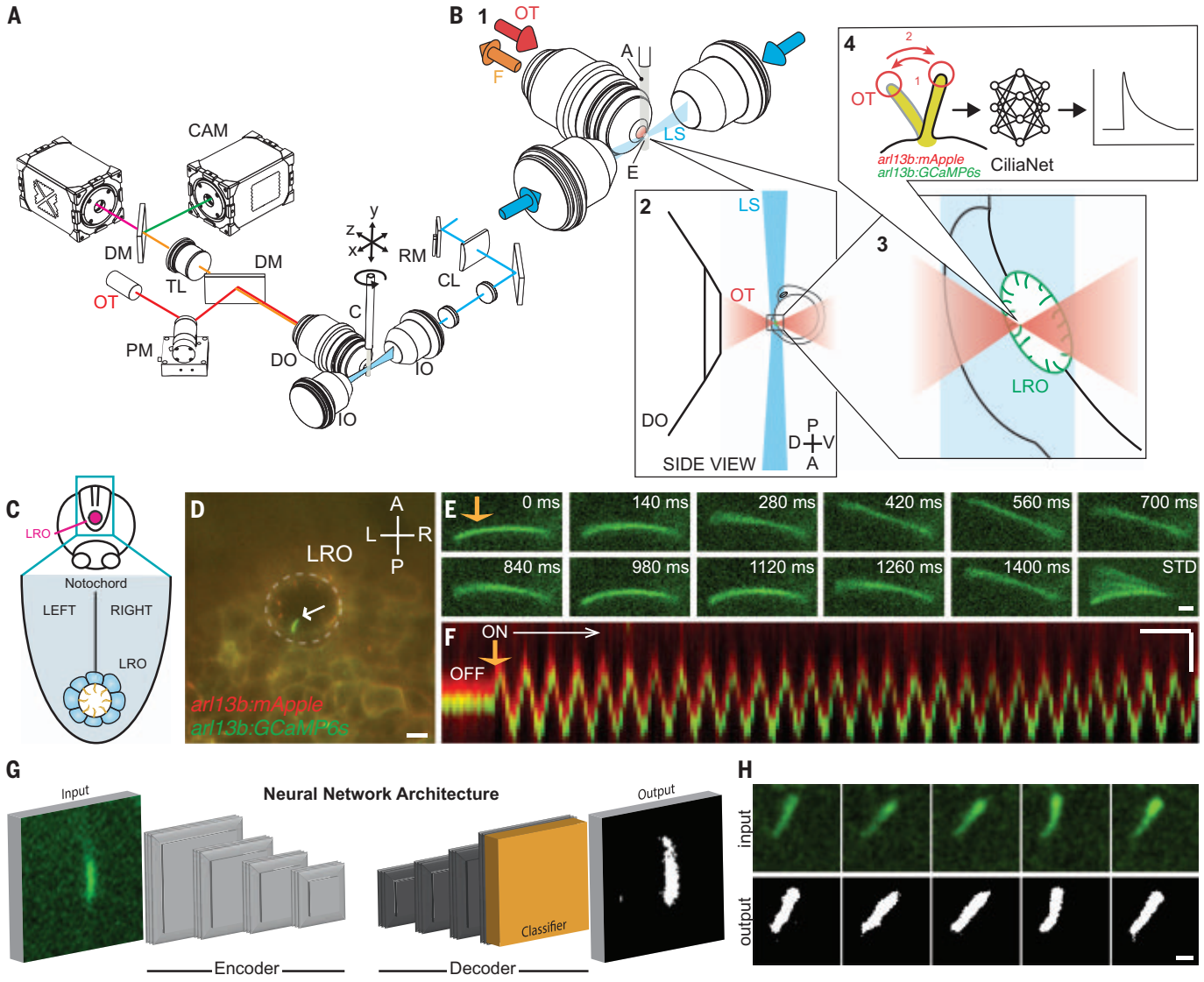


Fig. 1. CiliaSPOT is a precise and tunable platform for cilia mechanosensing studies. (**A**) Mechanical drawing of the CiliaSPOT microscope highlighting key components. For detailed description of the setup, see Materials and Methods and fig. S3. (**B**) Trapping and analyzing ciliary responses in the LRO of zebrafish embryos. (1) Close-up view showing the zebrafish embryo mounted in an agarose column extruded from a glass capillary. The cilia are fluorescently excited by the light sheet. (2) Side view of the optical trap and the LRO of a zebrafish embryo mounted in agarose. (3) Trapping of a single cilium in the LRO (green). (4) The trapped cilium is bent in an oscillatory fashion while being imaged. Images are then processed and analyzed by the CiliaNet machine learning algorithm to track and measure cilia responses to bending. A, agarose column; C, capillary; CL, cylindrical lens; CAM, camera; DM, dichroic mirror; DO, detection objective; E, embryo; F, fluorescence signal; IO, illumination objective; LS, light sheet; OT, optical trapping laser; PM, piezo mirror; RM, resonant mirror; TL, tube lens. (**C**) Illustration representing the LRO in the zebrafish embryo. (**D**) Representative image of an embryo expressing the ratiometric ciliary calcium reporter (*arl13b:mApple;arl13b:GCaMP6s*) with an intraciliary calcium oscillation (ICO, white arrow) within the LRO (dashed line). Scale: 10 μm. A, anterior; P, posterior; L, left; R, right; LRO, left-right organizer. (**E**) Representative montage of a fluorescent LRO cilium being bent by the optical trap (orange arrow) in vivo. Numbers indicate time after start of the bend. STD represents the standard deviation Z-projection of the montage. Scale: 2 μm. (**F**) Representative kymograph of an LRO cilium being bent by the optical trap (orange arrow). Note here that the GCaMP6s and mApple signals are kept slightly shifted for illustration purposes (see Materials and Methods and fig. S4). Scales: vertical: 2 μm; horizontal: 2 s. (**G**) Illustration of CiliaNet segmentation workflow. (**H**) Representative montage of sequential images of an LRO cilium dynamically trapped and moved by the optical tweezers (input) annotated by CiliaNet (output). Scale: 2 μm.

of LLC-PK1 cells with the fluorescent cilia marker transgene *arl13b:EGFP* (*20*) permitted us to validate simultaneous optical oscillatory bending and fluorescence microscopy (fig. S1, D to F, and movie S2).

The in vivo dynamics of immotile cilia in the zebrafish LRO were investigated with high-resolution differential interference contrast imaging (fig. S2 and movie S3). Measurements of ciliary angle and frequency displacement revealed that cilia on the left side of the LRO were subjected to greater angle displacements than cilia on the right side (fig. S2, B and C), consistent with the stronger flow on the left side of the LRO (*4*, *7*, *9*, *10*, *31*). We also observed small and slow displacement frequencies of LRO cilia (fig. S2D), with no significant difference between the two sides (fig. S2E). Our computer-controlled tweezer motions of LRO cilia were guided by these parameters, mimicking their in vivo bending dynamics.

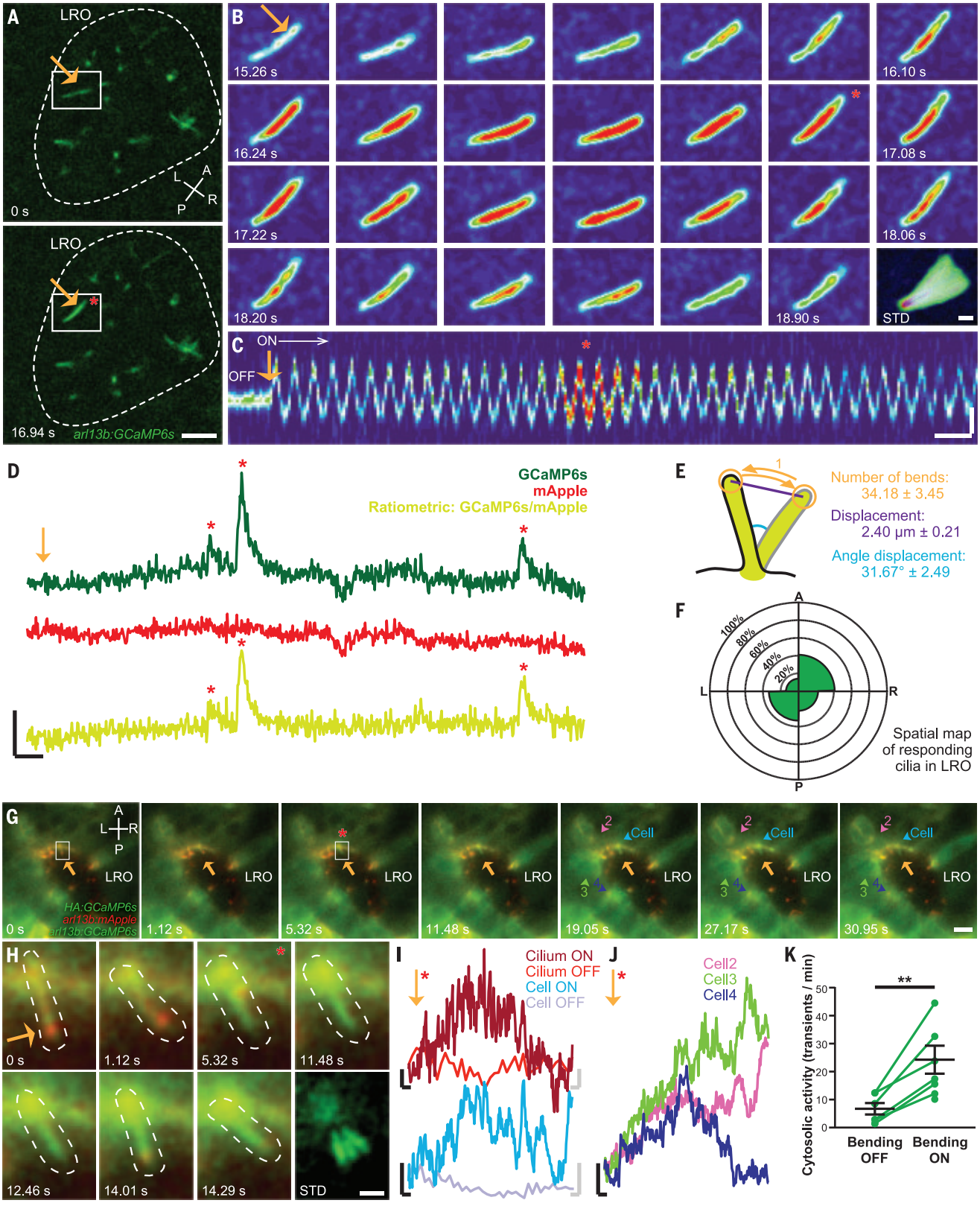


Fig. 2. Oscillatory mechanical stimuli on LRO cilia activate intraciliary calcium transients. (**A**) Representative images of the LRO of a *c21orf59* morphant zebrafish. Dashed line: LRO. Scale: 10 µm. (**B**) Representative montage of the GCaMP6s-positive LRO cilium highlighted in (A). Scale: 2 µm. (**C**) Kymograph of the cilium shown in (B) before ("OFF") and during optical bending ("ON", white arrow, starting at orange arrow). Scales: vertical: 2 µm; horizontal: 2 s. (**D**) Intraciliary intensity over time plots of a single LRO cilium exhibiting intraciliary calcium oscillations of different amplitudes in response to optical bending. Scales: vertical: 50% ΔF/F; horizontal: 5 s. (**E**) Optical bending characteristics associated with responding LRO cilia. Mean ± S.E.M, *n* = 23 responses analyzed. (**F**) Spatial

mapping of ciliary responses in the *c21orf59* embryos. Mean percentage of ciliary responses to optical bending in each region of the LRO (n = 88 cilia from 12 embryos). No statistical differences were observed between LRO regions (Fisher's exact tests with Bonferroni correction, all $P > 0.05$). (**G**) Representative montage of cytosolic calcium responses (colored arrowheads) following the intraciliary calcium response of a LRO cilium (white box) to oscillatory optical bending (orange arrow). Scale: 10 μm. (**H**) Representative montage of the cilium (dashed line) highlighted in (G). Scale: 2 μm. (**I**) GCaMP6s intensity over time plots of the cilium bent in (G) and (H) and in the connected cell (G), before (OFF) and after (ON) the start of the optical mechanical stimulation. Black scales for ON traces: vertical: 20% ΔF/F; horizontal: 2 s. Gray scales for OFF traces: vertical: 20% ΔF/F; horizontal: 0.3 s. (**J**) GCaMP6s intensity over time plots of the responding cells highlighted in (G). (**K**) Mean frequency of cytosolic activity (number of calcium transients per minute) at rest before bending (bending OFF) and during bending by the optical tweezers (bending ON) in *c21orf59* embryos (n = 6 morphants). $^{**}P < 0.01$ (P-value = 0.0044), paired two-tailed t-test. A, anterior; P, posterior; L, left; R, right; LRO, left-right organizer. Orange arrows: optical mechanical stimulation; red asterisks: start of the intraciliary response; numbers: time after start of the bend; STD: standard deviation Z-projection of montage.

To facilitate simultaneous live imaging of intraciliary calcium signaling and cilia deflection studies during LR development in zebrafish, we constructed a custom ciliary selective plane illumination microscope with optical tweezers ("CiliaSPOT", Fig. 1, A and B, and fig. S3). Selective plane illumination microscopy (light sheet microscopy) provides fast and gentle imaging of dynamic in vivo processes such as calcium signaling as a result of its rapid optical sectioning capabilities and reduced photobleaching (*32*, *33*).

The performance of CiliaSPOT was validated on immotile cilia in the LRO of 1-4 somite stage (ss) zebrafish embryos expressing a cilia-targeted ratiometric fluorescence calcium indicator system (*arl13b:GCaMP6s;arl13b:mApple)*, which we previously used to discover intraciliary calcium transients in the LRO (*20*). CiliaSPOT efficiently trapped the fluorescent cilia in the LRO with only 100 mW of laser power and was able to deflect immotile cilia in an oscillatory fashion mimicking their normal motions in vivo (Fig. 1, C to F, fig. S4, and movie S4). The amount of force applied to LRO cilia by CiliaSPOT was estimated to be 0.6 pN (fig. S5), in line with the estimated in vivo flow forces on LRO cilia (0.1 pN, Materials and Methods) and previous in vitro studies (*34*, *35*).

To analyze the fluorescent imaging data, we created the ciliary neural network ("CiliaNet") for the automated tracking of moving cilia and extraction of fluorescence signal changes (Fig. 1, G and H, fig. S6, and movie S5). CiliaNet permitted rapid analysis of our high-speed (7 Hz), two-channel recordings of ciliary calcium dynamics, even in large recordings of several thousands of frames, streamlining quantification of ciliary calcium responses to optical bending. We validated CiliaNet accuracy on CiliaSPOT recordings by comparing manual- and machine-analyzed datasets: CiliaNet was eight times faster than manual analysis with similar fidelity (fig. S6, B to D). Together, CiliaSPOT and CiliaNet provide a powerful platform for experimentally testing and analyzing ciliary mechanosensation and calcium signaling.

Mechanical stimulation of LRO cilia activates intraciliary calcium transients

Directional fluid flow in the LRO (described as leftward-biased in mice and counterclockwise in zebrafish), driven by motile cilia, is necessary for LR development (*3–5*, *7–10*, *12*, *13*). CiliaSPOT allowed us to determine the mechanistic link between this flow and the intraciliary calcium oscillations (ICOs) we and others have reported (*20*, *36*). To determine whether the ICOs are mechanically generated in LRO cilia, CiliaSPOT was used to deflect immotile LRO cilia in zebrafish without endogenous fluid flow. Knocking down *c21orf59* results in an absence of LRO flow, deficient intraciliary calcium signaling, and complete randomization of cardiac LR asymmetry (*20*, *37*), providing an ideal test setting. Any detectable elevations in intraciliary calcium or the presence of ICOs in *c21orf59* knockdown embryos must be solely due to our CiliaSPOT manipulation.

Intraciliary calcium transients were observed in immotile LRO cilia of *c21orf59* knockdown embryos trapped and bent in a controlled and oscillatory manner by CiliaSPOT (Fig. 2, A to D, Fig. 3, A to D, and movie S6). Prolonged mechanical oscillations, recapitulating the physiological ciliary behavior in the LRO of WT embryos with intact flow (*20*, *38*), triggered repetitive intraciliary calcium transients that resembled previously described ICOs (*20*, *36*) (Fig. 2D and movie S7). On average, intraciliary calcium transients occurred after 34.2 ± 3.5 bends, with a displacement of 2.4 ± 0.2 μm and a deflection angle of 31.7 ± 2.5° (Fig. 2E, mean ± SEM, fig. S7). All cilia regardless of their location within the LRO could respond to CiliaSPOT deflection with intraciliary calcium transients (Fig. 2F), suggesting that the endogenous left-sided LRO ICOs are generated in response to force exerted by directional LRO flow. Calcium transients in response to CiliaSPOT bending of LRO cilia were not spatially autonomous, as the calcium transients often spread to the cell body, neighboring LRO cells, and mesendodermal tissue around the LRO (Fig. 2, G to K, movie S8, and fig. S8).

CiliaSPOT oscillatory deflection of immotile LRO cilia of WT zebrafish (with intact endogenous counterclockwise fluid flow) activated intraciliary calcium transients (Fig. 3, F to J), regardless of the location of the manipulated cilia within the LRO (Fig. 3L). This not only reinforces our results from *c21orf59* knockdown embryos, but also suggests that immotile LRO cilia are capable of responding to mechanical stimuli, even while being subjected to intrinsic counterclockwise flow. Collectively, our results demonstrate that the cilium is a bona fide mechanosensor that mediates calcium signaling in the LRO.

The observed increases in calcium activity were not due to photodamage or excessive heat from CiliaSPOT. No calcium transients beyond the baseline were observed in the LRO of *c21orf59* knockdown embryos from the application of the 100 mW optical tweezers without oscillatory motion, indicating that CiliaSPOT did not induce aberrant calcium activity in the LRO (fig. S9 and movie S9). Further, we calibrated and employed a heat-sensitive fluorescent dye, Rhodamine B, as an optical thermometer to quantify potential temperature elevations in our optical trapping plane (*39*). The very small temperature rise from the 100 mW optical tweezers laser was minimal (~1°C, fig. S10) and thus unlikely to cause heat-associated damage. A digital thermometer, used to measure global temperature changes inside the specimen chamber of CiliaSPOT, found an imperceptible elevation from the 100 mW optical tweezers, within the margin of error of the thermometer (±0.1°C, Materials and Methods). Combined, these data strongly suggest the absence of any CiliaSPOT laser-illumination–associated side effects that could have contributed to a false positive calcium signal.

Ciliary mechanosensing at the LRO requires Polycystin-2

There are conflicting reports about whether the calcium-permeable cation channel Polycystin-2 (Pkd2) functions as a ciliary calcium channel and mechanosensor in a wide array of tissues (*5*, *14*, *23*, *24*, *36*, *40–42*). Pkd2 localizes to all LRO cilia and is critical for left-sided ICOs in the LRO (*20*, *36*) and normal LR patterning (*1*, *5*, *43–45*). In cultured ciliated renal cells, Pkd2 is proposed to be mechanosensitive for renal fluid flow, as application of artificial laminar fluid flow results in cytosolic calcium transients that require both the cilium and Pkd2 (*22*, *23*). Such studies seem consistent with Pkd2 functioning as a molecular mechanosensor on cilia, regulating intraciliary calcium in response to mechanical forces exerted by flow.

To resolve the function of Pkd2 in the LRO, we utilized CiliaSPOT to mechanically bend

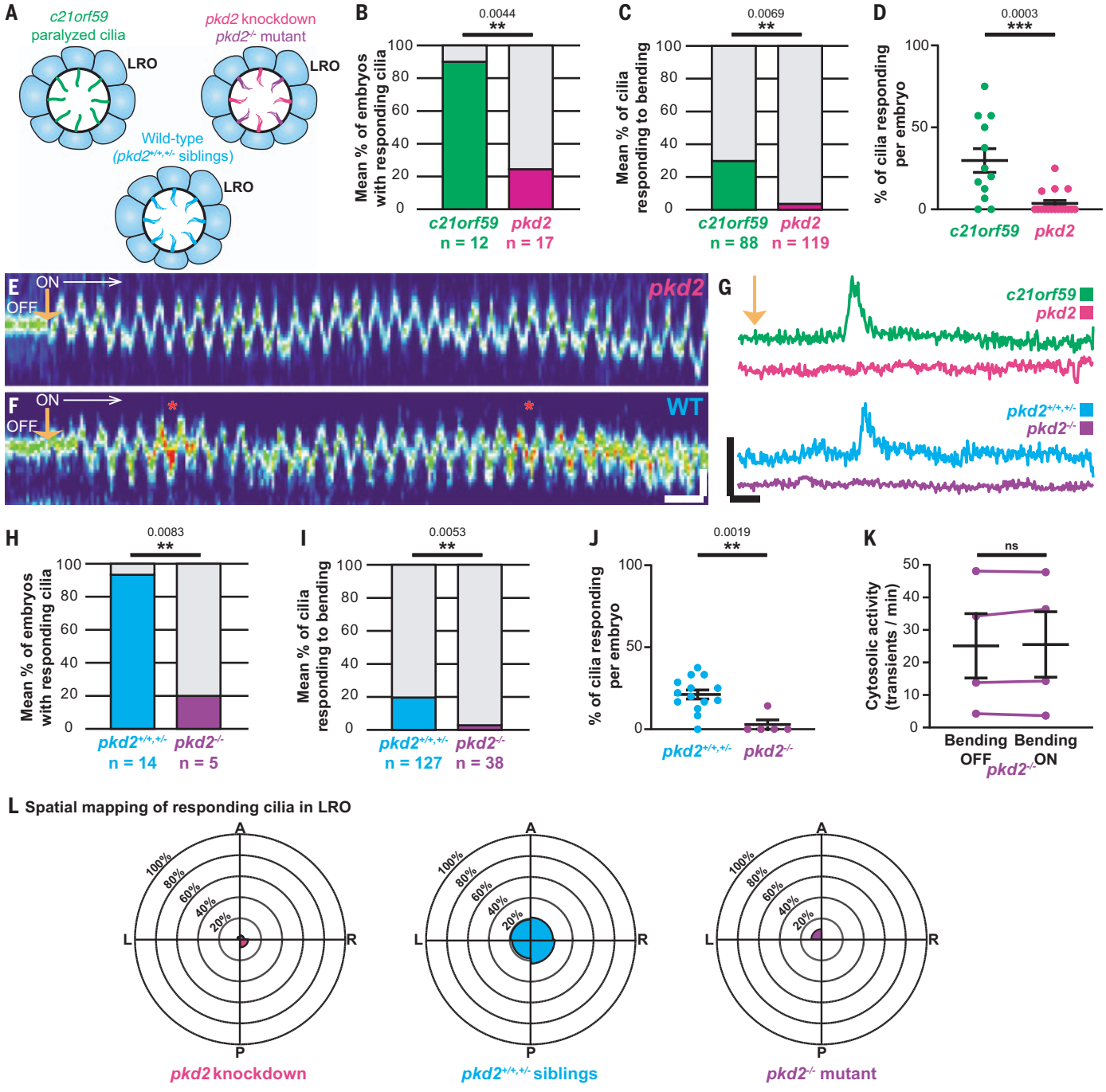


Fig. 3. Ciliary mechanosensation requires Polycystin-2. (**A**) Illustrations of the different models used in this study. (**B** to **D**) Percentage of embryos (B), cilia (C), and cilia per embryo (D) responding to optical bending in *c21orf59* (green, 88 cilia from 12 embryos) and *pkd2* morphants (magenta, 119 cilia from 17 embryos). Data shown are pooled from four independent experiments. Statistical comparison was analyzed by unpaired two-tailed *t*-tests; $^{**}P < 0.01$ and $^{***}P < 0.001$. (**E** and **F**) Representative kymographs of a LRO cilium from a *pkd2* morphant (E) and a WT (F) embryo showing their oscillatory motions and calcium activity in response to the optical bending (ON, white arrow, starting at orange arrow). Scales: vertical: 2 μm; horizontal: 2 s. (**G**) Representative GCaMP6s intraciliary intensity over time plots of a single LRO cilium in response to optical bending in *c21orf59* (green) and *pkd2* (magenta) morphants, and in a *pkd2* mutant (purple) and WT sibling (blue). Scales: vertical: 100% ΔF/F; horizontal: 5 s. (**H** to **J**) Percentage of embryos (H), cilia (I), and cilia per embryo (J) responding to optical bending in WT siblings (total of 127 cilia from 14 embryos) and *pkd2* homozygous mutants (total of 38 cilia from 5 embryos). Data shown are pooled from five independent experiments. Statistical comparison was analyzed by unpaired two-tailed *t*-tests; $^{**}P < 0.01$. (**K**) Mean frequency of cytosolic activity (number of calcium transients per minute) at rest before bending (bending OFF) compared with when LRO cilia are being bent by the optical tweezers (bending ON) in *pkd2* mutants. $n = 4$ mutants. ns, not significant, paired two-tailed *t*-test. (**L**) Spatial mapping of ciliary responses to optical bending in the *pkd2* knockdown (magenta), *pkd2* homozygous mutant (purple) and WT sibling ($pkd2^{+/+;+/-}$, blue) zebrafish LROs. The rose diagrams represent the mean percentage of ciliary responses to optical bending in each region of the LRO. (*pkd2* knockdown = 119 cilia from 17 embryos; *pkd2* homozygous mutants = 38 cilia from 5 embryos; WT siblings $pkd2^{+/+;+/-}$ = 127 cilia from 14 embryos). A, anterior; P, posterior; L, left; R, right.

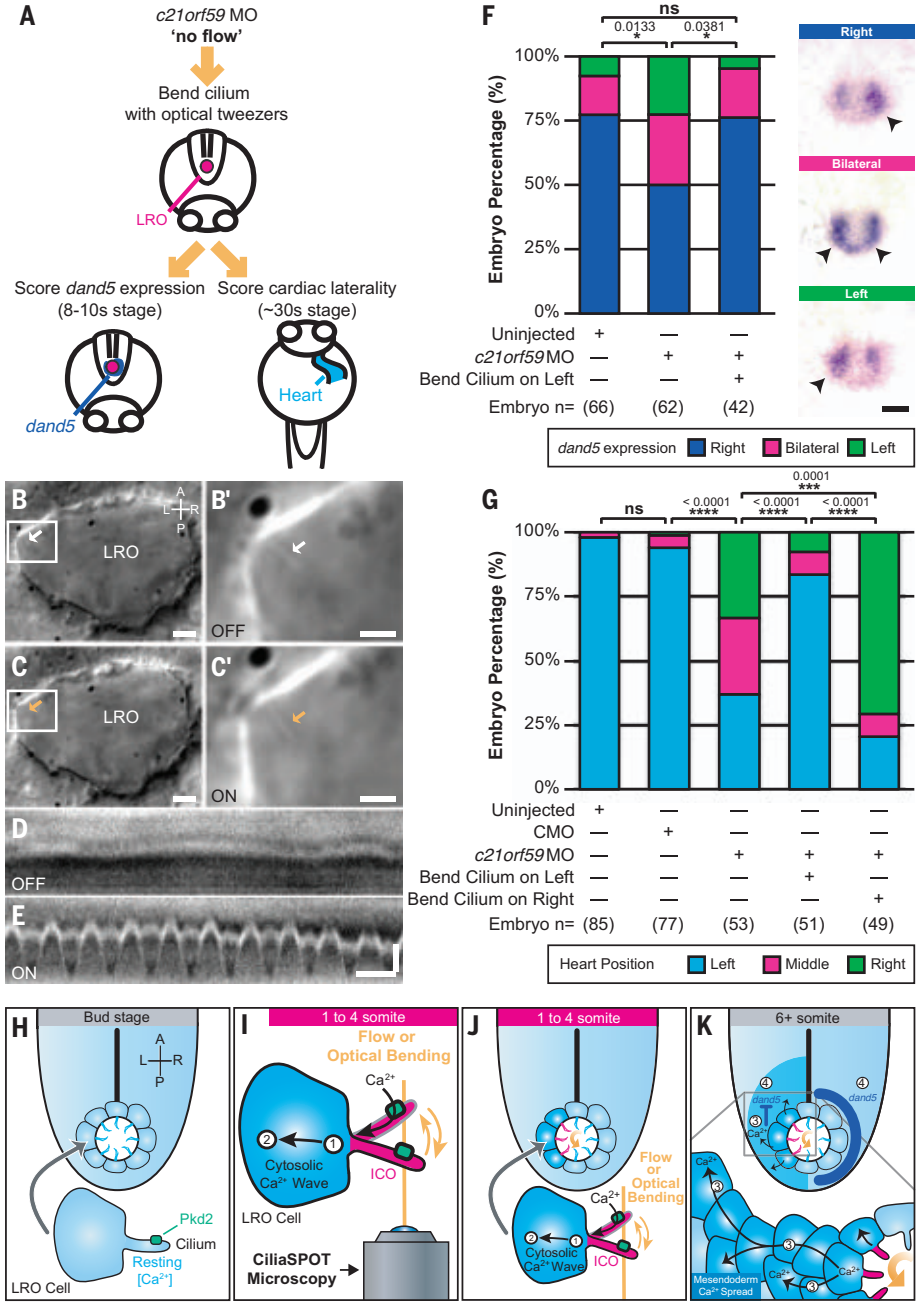


Fig. 4. Ciliary mechanosensation is determinative for LR asymmetry. (**A**) Schematic depicting the approach followed to assess *dand5* expression and cardiac laterality in *c21orf59* morphants (MO) after oscillatory optical bending of one LRO cilium. (**B** to **C'**) Representative images of LRO cilia (arrows) from *c21orf59* embryos in the absence [(B) white arrow, OFF] or presence of optical tweezers [(C) orange arrow, ON]. [(B') and (C')] Cilium highlighted in white box in [(B) and (C)]. Scales: [(B) and (C)]: 5 μm; [(B') and (C')]: 2 μm. (**D** and **E**) Representative kymographs of LRO cilia from *c21orf59* embryos in the absence [(D) OFF] or presence of oscillatory optical tweezers [(E) ON]. Scales: vertical: 1 μm; horizontal: 1 s. (**F**) Graph with illustrative pictures, representing percentage of uninjected and *c21orf59* morphants displaying normal right-sided (dark blue) and abnormal left-sided (green) or bilateral (magenta) *dand5* expression. *n* = total number of embryos analyzed. Statistical comparison was analyzed by a Pearson's chi-square test (Bonferroni corrected); $*P < 0.05$ and ns: $P \geq 0.05$. Scale: 50 μm. (**G**) Graph represents percentage of uninjected, control morpholino-injected (CMO) and *c21orf59* morphants displaying normal left-sided (light blue) and abnormal right-sided (green) or middle (magenta) positioned hearts. *n* = total number of embryos analyzed. Data shown are pooled from three independent experiments. Statistical comparison was analyzed by one-way ANOVA with Tukey's multiple comparison test; $***P < 0.001$, $****P < 0.0001$ and ns: $P \geq 0.05$. (**H** to **K**) Model for calcium-mediated ciliary mechanosensation in the LRO during LR development. At early stages of LR patterning, counterclockwise left-biased flow (curved orange arrow) or ciliary optical bending triggers Pkd2-dependent intraciliary calcium signaling (in cilia, magenta ICOs; in cells, dark blue) on the side of the LRO subjected to ciliary mechanical stimulation (1). Cilia-to-cytosolic LRO calcium (2) is then transmitted to neighboring cells of the mesendoderm in a side-biased manner (3), which in turn ultimately direct asymmetric gene expression (4) leading to LR patterning. A, anterior; P, posterior; L, left; R, right.

cilia in the LRO of *pkd2* mutant and knockdown zebrafish embryos, and measured intraciliary calcium levels. Notably, we found a significantly lower incidence of CiliaSPOT-induced intraciliary calcium transients in immotile LRO cilia of both *pkd2* knockdown and *pkd2* knockout embryos (Fig. 3 and movie S10). Further, there was no elevation in cytosolic calcium activity after LRO ciliary bending in *pkd2* mutant embryos (Fig. 3K). This loss of intraciliary calcium transients in optically tweezed *pkd2* mutant embryos confirms that the CiliaSPOT-induced intraciliary calcium transients in WT and *c21orf59* knockdown embryos were not caused by photodamage or microdamage. These results support a role for Pkd2 as a mechanosensitive calcium channel on cilia in the LRO.

Ciliary mechanosensation is instructive for LR asymmetry

Directional LRO fluid flow is essential for asymmetric gene expression of Nodal signaling components during LR development in mice, zebrafish, and *Xenopus* (*2*, *4*, *7*–*10*). We and others previously linked directional fluid flow and ICOs to left-sided degradation of *dand5*, the first asymmetrically expressed gene and an upstream inhibitor of Nodal signaling (*7*, *46*, *47*). To test whether ciliary mechanosensing itself is sufficient for asymmetric *dand5* expression, we utilized optical tweezers to bend cilia in *c21orf59* knockdown zebrafish, which have LR-randomized expression of *dand5* in the LRO (*20*). Oscillatory deflection of one immotile cilium was performed at the 1 somite stage (1 ss) on the left side of the LRO of *c21orf59* knockdown embryos for 1 hour, within the critical window when initial symmetry breaking occurs in the zebrafish LRO (1-4 ss) (*20*) (Fig. 4, A to E). After tweezing, embryos were unmounted from the microscope, raised normally until 8-10 ss, and assayed for asymmetric *dand5* expression by in situ hybridization (*46*, *47*). Strikingly, mechanical stimulation of a single

cilium on the left side of the LRO was sufficient to rescue proper right-biased *dand5* expression (Fig. 4F).

Ciliary bending can rescue the laterality of asymmetric organs such as the heart, similar to experiments in the mouse LRO utilizing artificial directional fluid flow (*3*). We optically bent cilia in *c21orf59* knockdown zebrafish, which normally exhibit complete LR randomization of cardiac jogging, a precursor for asymmetric cardiac looping in zebrafish (illustrated in movie S11) (*20*, *37*). After tweezer stimulation of a single immotile cilium on the left side of the LRO for 1 hour, embryos were raised until 24 hours post fertilization (hpf) and scored for the LR direction of cardiac jogging. Bending a cilium on the left side of the LRO resulted in normal, leftward cardiac jogging (Fig. 4G), similar to the rescue of asymmetric *dand5* expression described above. These rescue data further confirmed that our optical tweezers approach did not negatively affect LR development or embryonic viability. Bending a cilium on the right side of the LRO led to the opposite—rightward cardiac jogging (Fig. 4G). These results indicate that ciliary mechanosensation is sufficient and instructive for both LR asymmetric gene expression and cardiac situs.

Finally, we examined whether Pkd2 is required for ciliary mechanosensing during cardiac LR patterning by applying the same approach to *pkd2* knockdown embryos, which also have LR randomized cardiac looping and jogging (*43*, *44*). In contrast to *c21orf59* embryos, bending immotile cilia on the left side of the LRO for 1 hour in *pkd2* knockdown embryos failed to rescue proper leftward cardiac jogging (fig. S11), demonstrating the key role of Pkd2 as a mechanosensor required in LRO cilia for cardiac situs determination. Altogether, our findings establish ciliary force sensing as necessary, sufficient, and instructive for LR asymmetry.

Discussion

The importance of flow in the LRO has long been realized and a variety of experimental investigations have shown that its absence leads to aberrant LR asymmetry of molecular markers and organ development (*14*, *15*, *17*). A large body of data has accumulated that are consistent with flow mediating these LR differences by the transport of morphogens or by mechanosensation, but without clear demonstration of either. Our results directly show that LRO ciliary mechanosensation is both necessary and sufficient for molecular and morphological LR asymmetry.

Using carefully controlled optomechanical manipulations of cilia within the LRO, we provide evidence that Pkd2-dependent intraciliary calcium transients are mechanically generated in the LRO and that LRO ciliary mechanosensation is determinative for LR development (Fig. 4, H to K). By applying an oscillatory bending force directly to cilia for up to an hour, we emulated physiological conditions for in vivo cilia deflection at the LRO and demonstrated that oscillatory bending on a scale of seconds to minutes is critical to stimulate a calcium response in LRO cilia. This similarity to endogenous LRO ICOs (*20*, *36*), that occur at a low frequency (ranging from 3.3 to 5.2 min) (*20*), indicates that a single cilium requires repetitive and consistent mechanical stimulation before a calcium signal is transmitted to the cell. This allows the cilium to discriminate between true (repetitive) and false (random) stimuli by filtering out extraneous biological noise in tissues subjected to abundant biomechanical forces, including directional extracellular fluid flow and nondirectional cytoskeletal movements. We posit that the cilium's noise filtering capability may be adaptable, as the filter threshold in WT embryos with intrinsic flow appears to be higher than in flowless embryos as suggested by the lower number of responses in WT LRO cilia (Fig. 3, I and J) compared with *c21orf59* cilia (Fig. 3, C and D).

Our results suggest that LRO cilia may be heterogeneous, as among all the LRO cilia we optically tweezed, 20 to 30% responded positively to the oscillatory optical bending (Fig. 3, C and I). Notably, our previous study found the same percentage of cilia that display ICOs in WT LROs with endogenous flow (*20*). These observations, coupled with another report that found a similar percentage of immotile cilia in the LRO (*31*), suggest that a distinct subpopulation of LRO cilia function as mechanosensors as hypothesized in the two-cilia model of LR development (*1*, *19*). Although our findings experimentally validate this model, we cannot definitively rule out that a morphogen participates in LR development (*2*, *18*). It remains possible that mechanosensation and chemosensation cooperate at the LRO; for example, mechanical cues might expose ligand receptor sites on cilia or Pkd2 itself.

Investigating endogenous ciliary dynamics in WT LROs revealed that left-sided LRO cilia experienced a greater angle displacement than right-sided ones (fig. S2B). We found that the 75th percentile of WT cilia on the left side of the LRO exhibited an angular displacement of 31.2 ± 1.5° (mean ± SEM, fig. S2F), which mirrored the mean angular displacement of tweezed cilia that responded positively to optical bending (31.7 ± 2.5°; Fig. 2E). These results suggest that a greater displacement angle due to force exerted by directional flow increases the likelihood of an intraciliary calcium transient on the left-side of the LRO.

Our conclusions in the zebrafish LRO utilizing CiliaSPOT mirrored results in the mouse LRO in an accompanying research article (*48*). Both studies independently interrogated the function of cilia as calcium-mediated mechanosensors in the LRO using different experimental models and cilia deflection approaches. Despite this, we found surprising similarities in ciliary mechanosensation and intraciliary calcium transients between the mouse and zebrafish LRO, emphasizing an evolutionarily conserved mechanism for cilia mechanotransduction during LR development in vertebrates with ciliated LROs. Intriguing findings from previous work in mice and zebrafish point to slow, localized LRO flow currents as sufficient for normal LR patterning (*9*, *49*), suggesting that fast, coordinated LRO flow at later stages is not essential for initiating asymmetry. Our work demonstrates that mechanical stimulation of a single cilium in the LRO for extended periods of time was determinative for molecular and cardiac asymmetry, resolving prior discrepancies (*24*). These findings show that ciliary mechanosensing is a fundamental and potent cellular signaling mechanism in development and disease, as a single ciliated cell was sufficient to create chiral vertebrate asymmetry. Finally, our study reveals that, in addition to biochemical and molecular mechanisms, small physical forces at the level of single cells and organelles play essential roles in shaping the bilateral body of the developing embryo.

REFERENCES AND NOTES

1. J. McGrath, S. Somlo, S. Makova, X. Tian, M. Brueckner, *Cell* **114**, 61–73 (2003).
2. Y. Okada *et al*., *Mol. Cell* **4**, 459–468 (1999).
3. S. Nonaka, H. Shiratori, Y. Saijoh, H. Hamada, *Nature* **418**, 96–99 (2002).
4. S. Nonaka *et al*., *Cell* **95**, 829–837 (1998).
5. S. Yoshiba *et al*., *Science* **338**, 226–231 (2012).
6. M. Levin, R. L. Johnson, C. D. Stern, M. Kuehn, C. Tabin, *Cell* **82**, 803–814 (1995).
7. A. Schweickert *et al*., *Curr. Biol.* **20**, 738–743 (2010).
8. J. J. Essner, J. D. Amack, M. K. Nyholm, E. B. Harris, H. J. Yost, *Development* **132**, 1247–1260 (2005).
9. P. Sampaio *et al*., *Dev. Cell* **29**, 716–728 (2014).
10. N. Okabe, B. Xu, R. D. Burdine, *Dev. Dyn.* **237**, 3602–3612 (2008).
11. J. J. Essner *et al*., *Nature* **418**, 37–38 (2002).
12. A. G. Kramer-Zucker *et al*., *Development* **132**, 1907–1921 (2005).
13. W. Supatto, S. E. Fraser, J. Vermot, *Biophys. J.* **95**, L29–L31 (2008).
14. L. Djenoune, K. Berg, M. Brueckner, S. Yuan, *Nat. Rev. Cardiol.* **19**, 211–227 (2022).
15. R. B. Little, D. P. Norris, *Semin. Cell Dev. Biol.* **110**, 11–18 (2021).
16. R. R. Ferreira, A. Vilfan, F. Jülicher, W. Supatto, J. Vermot, *eLife* **6**, e25078 (2017).
17. R. R. Ferreira, H. Fukui, R. Chow, A. Vilfan, J. Vermot, *J. Cell Sci.* **132**, jcs213496 (2019).
18. Y. Tanaka, Y. Okada, N. Hirokawa, *Nature* **435**, 172–177 (2005).
19. C. J. Tabin, K. J. Vogan, *Genes Dev.* **17**, 1–6 (2003).
20. S. Yuan, L. Zhao, M. Brueckner, Z. Sun, *Curr. Biol.* **25**, 556–567 (2015).
21. D. Takao *et al*., *Dev. Biol.* **376**, 23–30 (2013).
22. H. A. Praetorius, K. R. Spring, *J. Membr. Biol.* **184**, 71–79 (2001).
23. S. M. Nauli *et al*., *Nat. Genet.* **33**, 129–137 (2003).
24. M. Delling *et al*., *Nature* **531**, 656–660 (2016).
25. S. Su *et al*., *Nat. Methods* **10**, 1105–1107 (2013).
26. A. Ashkin, K. Schütze, J. M. Dziedzic, U. Euteneuer, M. Schliwa, *Nature* **348**, 346–348 (1990).
27. A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, S. Chu, *Opt. Lett.* **11**, 288 (1986).
28. K. Svoboda, S. M. Block, *Annu. Rev. Biophys. Biomol. Struct.* **23**, 247–285 (1994).
29. A. K. O'Connor *et al*., *Cilia* **2**, 8 (2013).
30. J. G. Goetz *et al*., *Cell Rep.* **6**, 799–808 (2014).
31. B. Tavares *et al*., *eLife* **6**, e25165 (2017).
32. J. Huisken, D. Y. Stainier, *Opt. Lett.* **32**, 2608–2610 (2007).
33. T. V. Truong, W. Supatto, D. S. Koos, J. M. Choi, S. E. Fraser, *Nat. Methods* **8**, 757–760 (2011).

34. C. Battle, C. M. Ott, D. T. Burnette, J. Lippincott-Schwartz, C. F. Schmidt, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1410–1415 (2015).
35. A. Resnick, *J. Biomed. Opt.* **15**, 015005 (2010).
36. K. Mizuno *et al*., *Sci. Adv.* **6**, eaba1195 (2020).
37. C. Austin-Tse *et al*., *Am. J. Hum. Genet.* **93**, 672–686 (2013).
38. M. T. Boskovski *et al*., *Nature* **504**, 456–459 (2013).
39. D. Moreau, C. Lefort, R. Burke, P. Leveque, R. P. O'Connor, *Biomed. Opt. Express* **6**, 4105–4117 (2015).
40. P. G. DeCaen, M. Delling, T. N. Vien, D. E. Clapham, *Nature* **504**, 315–318 (2013).
41. M. Delling, P. G. DeCaen, J. F. Doerner, S. Febvay, D. E. Clapham, *Nature* **504**, 311–314 (2013).
42. S. J. Kleene, N. K. Kleene, *Am. J. Physiol. Renal Physiol.* **320**, F1165–F1173 (2021).
43. P. Pennekamp *et al*., *Curr. Biol.* **12**, 938–943 (2002).
44. B. W. Bisgrove, B. S. Snarr, A. Emrazian, H. J. Yost, *Dev. Biol.* **287**, 274–288 (2005).
45. J. Schottenfeld, J. Sullivan-Brown, R. D. Burdine, *Development* **134**, 1605–1615 (2007).
46. S. S. Lopes *et al*., *Development* **137**, 3625–3632 (2010).
47. H. Hashimoto *et al*., *Development* **131**, 1741–1753 (2004).
48. T. A. Katoh *et al*., *Science.* **379**, 66–71 (2023).
49. K. Shinohara *et al*., *Nat. Commun.* **3**, 622 (2012).
50. L. Djenoune *et al*., CiliaNet: a deep learning tool for cilia analysis, v.10, Zenodo (2022); https://doi.org/10.5281/zenodo.7417672

ACKNOWLEDGMENTS

We thank M. Khokha for insightful discussions; H. Shroff, J. Huisken, R. Power, J. Choi, and M. Guo for providing light sheet microscopy advice; D. Sosnovik for helpful suggestions regarding optical thermometer experiments; Z. Sun, L. Trinh, and Y. Li for advice on zebrafish experiments and sharing lines; S. Somlo, Y. Cai, and S. Makova for sharing cell lines; M. Khokha, S. Makova, and Z. Sun for critical feedback on the manuscript; and A. Brugger, N. Djenoune, and M. Jones for technical assistance. **Funding:** This work was supported by the American Heart Association (Career Development Award 940516 to S.Y., Postdoctoral Fellowship Award 830304 to L.D., Transformational Project Award 969048 to S.Y.), the Charles Hood Foundation (Child Health Award to S.Y.), the Gordon and Betty Moore Foundation (grant 3396 to S.E.F.), the Hassenfeld Foundation (Scholar Award to S.Y.), Massachusetts General Hospital (Institutional funds from Department of Medicine and Cardiology Division to S.Y.), the National Institutes of Health (Outstanding Investigator Award 1R35HL145249 to M.B., Pathway to Independence Award 1K99HD086274 to S.Y., grant 1R01HL151704 to C.T.N., grant 1R01HL165241 to S.Y.), the National Science Foundation (grant 1608744 to S.E.F.), University of Southern California Translational Imaging Center (S.E.F. and T.V.T.), and Yale University (Institutional funds from Departments of Biochemistry and Biophysics to J.H., and Department of Pediatrics to S.Y.). **Author contributions:** Conceptualization: L.D., M.M., M.B., and S.Y. Methodology: L.D., M.M., T.V.T., C.T.N., S.E.F., J.H., and S.Y. Investigation: L.D., M.M., and S.Y. Funding acquisition: L.D., C.T.N., S.E.F., M.B., J.H., and S.Y. Project administration: S.E.F., M.B., J.H., and S.Y. Supervision: S.E.F., M.B., J.H., and S.Y. Writing – original draft: L.D., M.M., and S.Y. Writing – review and editing: all authors. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data are available in the manuscript or supplementary materials. Zebrafish lines and plasmids are available upon request. CiliaNet is available on Github at https://github.com/shiaulouyuan/CiliaNet and archived on Zenodo (*50*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq7317
Materials and Methods
Figs. S1 to S11
Tables S1 and S2
References (*51*–*73*)
MDAR Reproducibility Checklist
Movies S1 to S11

Submitted 9 May 2022; resubmitted 3 November 2022
Accepted 9 December 2022
10.1126/science.abq7317

GLACIERS

Global glacier change in the 21st century: Every increase in temperature matters

David R. Rounce[1,2]*, Regine Hock[2,3], Fabien Maussion[4], Romain Hugonnet[5,6,7], William Kochtitzky[8,9], Matthias Huss[5,6,10], Etienne Berthier[7], Douglas Brinkerhoff[11], Loris Compagno[5,6], Luke Copland[8], Daniel Farinotti[5,6], Brian Menounos[12,13], Robert W. McNabb[14]

Glacier mass loss affects sea level rise, water resources, and natural hazards. We present global glacier projections, excluding the ice sheets, for shared socioeconomic pathways calibrated with data for each glacier. Glaciers are projected to lose 26 ± 6% (+1.5°C) to 41 ± 11% (+4°C) of their mass by 2100, relative to 2015, for global temperature change scenarios. This corresponds to 90 ± 26 to 154 ± 44 millimeters sea level equivalent and will cause 49 ± 9 to 83 ± 7% of glaciers to disappear. Mass loss is linearly related to temperature increase and thus reductions in temperature increase reduce mass loss. Based on climate pledges from the Conference of the Parties (COP26), global mean temperature is projected to increase by +2.7°C, which would lead to a sea level contribution of 115 ± 40 millimeters and cause widespread deglaciation in most mid-latitude regions by 2100.

Glaciers, here referring to all glacial land ice excluding the Greenland and Antarctic ice sheets, are responsible for 21 ± 3% of sea level rise from 2000 to 2019, contributing 0.74 ± 0.04 mm sea level equivalent (SLE) yr^{-1} (*1*). Projections suggest this contribution could increase to 2.5 mm SLE yr^{-1} by 2100 (*2*). Glaciers are also a critical water resource for ~1.9 billion people (*3*), and projected losses will alter water availability impacting annual and seasonal runoff (*4*). Glacier-related hazards, including glacier outburst floods, are also expected to change in frequency and magnitude over the next century as a result of mass loss (*5*). Projecting the magnitude, spatial pattern, and timing of glacier mass loss is therefore essential to support climate adaptation and mitigation efforts for communities ranging from the coast to the high mountains.

Previous projections of glacier mass loss from the glacier model intercomparison project (GlacierMIP) (*2*) estimated glacier contribution to sea level rise for ensembles of representative concentration pathways (RCPs), and results were extended to shared socioeconomic pathways (SSPs) using statistical models of these simulations (*6*). GlacierMIP provided these projections at regional scales based on simulations from 11 glacier evolution models that varied with respect to the complexity of model physics, simulated physical processes, model calibration, spatial resolution, and modeling domain. Calibration data varied from in situ measurements of less than 300 of the world's more than 215,000 glaciers to regional geodetic and/or gravimetric mass balance observations. Furthermore, only one global model simulated glacier dynamics using a flowline model (*7*), whereas all others relied on empirical volume-area scaling or parameterizations of mass redistribution; only one model accounted for frontal ablation (i.e., the sum of iceberg calving and submarine melt) of marine-terminating glaciers (*8*), whereas all others treated any glacier as land-terminating; further, no global model accounted for debris cover. Existing multimodel projections (*2*, *6*, *9*) are thus limited to regional scales and neglect key physical processes controlling glacier mass loss.

We produce a set of global glacier projections for every glacier on Earth for SSPs from 2015 to 2100 by leveraging global glacier mass balance data (*1*) and near-global frontal ablation data (*10*–*13*). To provide policy-relevant scenarios, our projections are grouped based on mean global temperature increases by the end of the 21st century compared with preindustrial levels to explicitly link differences in glacier mass loss, sea level rise, and the number of glaciers that vanish in response to changes in mean global temperature. Our glacier evolution model, a hybrid of the Python Glacier Evolution Model (PyGEM) (*14*, *15*) and Open Global Glacier Model (OGGM) (*7*), enables us to produce global glacier projections that explicitly account for glacier dynamics using a

[1]Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. [2]Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK, USA. [3]Department of Geosciences, University of Oslo, Oslo, Norway. [4]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria. [5]Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH Zürich, Zurich, Switzerland. [6]Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland. [7]LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France. [8]Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, Ontario, Canada. [9]School of Marine and Environmental Programs, University of New England, Biddeford, ME, USA. [10]Department of Geosciences, University of Fribourg, Fribourg, Switzerland. [11]Department of Computer Science, University of Montana, Missoula, MT, USA. [12]Geography Earth and Environmental Sciences, University of Northern British Columbia, Prince George, BC, Canada. [13]Hakai Institute, Campbell River, BC, Canada. [14]School of Geography and Environmental Sciences, Ulster University, Coleraine, UK.
***Corresponding author. Email: drounce@cmu.edu**

flowline model (*7*) based on the shallow-ice approximation (*16*), the effects of debris thickness on sub-debris melt rates (*17*), and frontal ablation (*8*). Our estimates of glacier contribution to sea level rise also account for the ~15% of ice from marine-terminating glaciers that is already below sea level (*18*). Projections are also reported for SSPs and RCPs to highlight differences compared with previous studies.

Projections of policy-relevant scenarios

The Paris Agreement, adopted in 2015 by 195 countries, agreed to keep the increase in global mean temperature by the end of the 21st century relative to preindustrial levels below 2°C, and that efforts should be made to limit the temperature change to 1.5°C. This target was kept alive in the Glasgow Agreement adopted by the Conference of the Parties (COP26) in 2021. To evaluate the sensitivity of glaciers to global mean temperature increases, the glacier projections are aggregated into +1.5°C, +2°C, +3°C, and +4°C temperature change scenarios by 2100 relative to preindustrial levels (Fig. 1).

Globally, glaciers are projected to lose 26 ± 6% (+1.5°C) to 41 ± 11% (+4°C) of their mass by 2100, relative to 2015 [ensemble median ± 95% confidence interval (CI)]. This mass loss would increase mean sea level by 90 ± 26 mm SLE under the +1.5°C scenario and 99 ± 31 mm SLE under the +2°C scenario. The higher temperature change scenarios of +3°C and +4°C lead to contributions of 125 ± 39 and 154 ± 44 mm SLE, respectively, highlighting a 71% increase between the +1.5°C and +4°C scenarios.

The rate of sea level rise from glacier mass loss near the end of the 21st century ranges from 0.70 ± 0.45 to 2.23 ± 1.08 mm SLE yr^{-1} depending on the temperature change scenario (fig. S1). For +1.5°C, the rate of sea level rise peaks at 1.29 ± 0.59 mm SLE yr^{-1} around 2035 and declines thereafter whereas the rate for +4°C steadily increases for the remainder of this century. Similar trends are observed in the area-averaged mass loss rate, where the maximum loss rate of 0.82 ± 0.36 m water equivalent (w.e.) yr^{-1} occurs around 2035 before diminishing to 0.59 ± 0.34 m w.e. yr^{-1} at the end of the century for the +1.5°C scenario; the mass loss rate continuously increases to 2.02 ± 1.30 m w.e. yr^{-1} by the end of the century for the +4°C scenario (Fig. 1I). Even if the global mean temperature change is limited to +1.5°C, we estimate that 104,000 ± 20,000 glaciers (49 ± 9% of the total inventoried) will disappear by 2100 and at least half of those will be lost before 2050 (Fig. 1E). Most of the glaciers projected to disappear are <1 km^2 (Fig. 2) but regardless of their small size, their disappearance may still negatively affect local hydrology, tourism, glacier hazards, and cultural values (*19*). Glaciers projected to disappear represent 2 to 8% of the glacier contribution to sea level rise depending on the temperature change scenario.

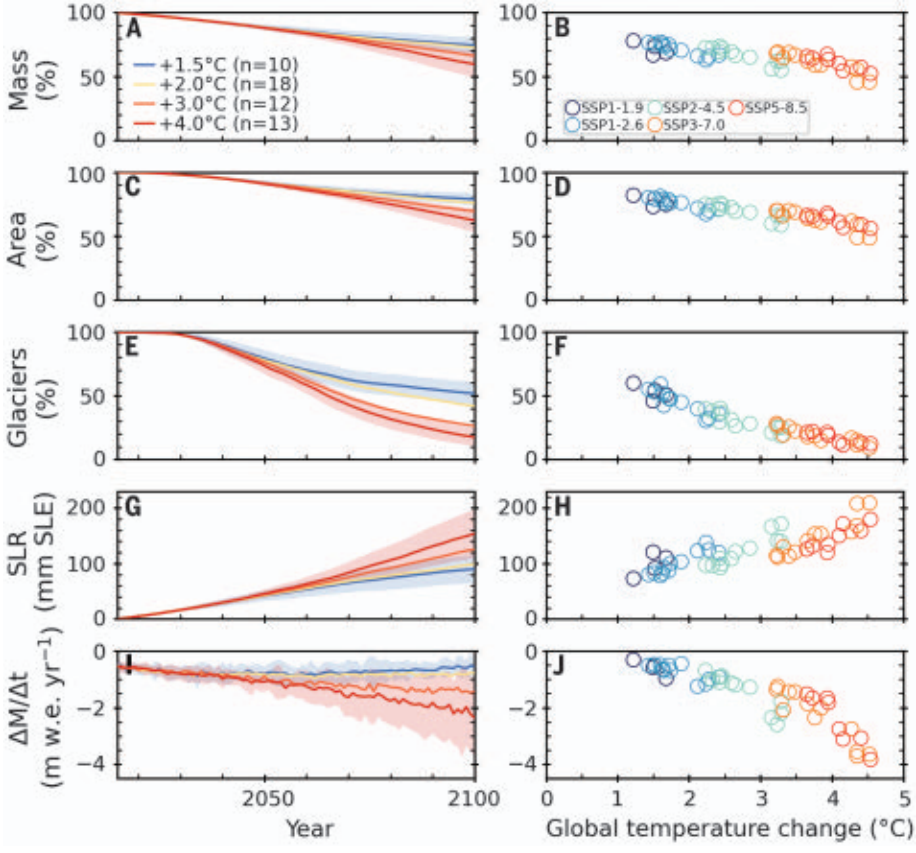


Fig. 1. Projected global glacier changes for scenarios of global mean temperature change. (**A** and **B**) Mass remaining, (**C** and **D**) area remaining, (**E** and **F**) glaciers remaining, (**G** and **H**) sea level rise (SLR) contributed from glaciers, and (**I** and **J**) area-averaged mass change rate for all glaciers globally. Projections are shown from 2015 to 2100 (left panels), and at 2100 (right panels). Values in [(A) to (H)] are relative to 2015. Colors depict the global mean temperature change scenarios (left panels) and the SSPs corresponding to the global temperature changes (right panels). The number (*n*) of glacier projections with different general circulation models (GCMs) and SSPs that fall into each temperature change scenario is shown in the legend. Lines (left panels) show the ensemble median and shading indicates the 95% CI for each temperature change scenario.

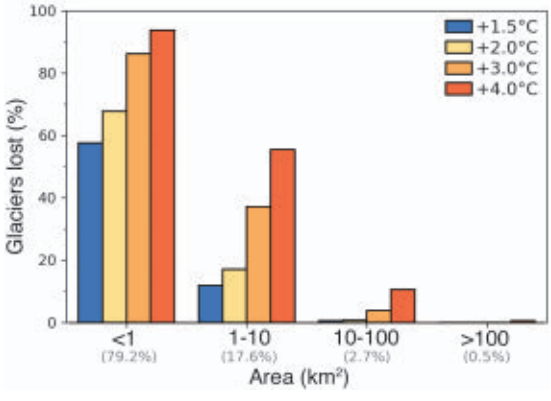


Fig. 2. Percent of glaciers projected to vanish between 2015 and 2100 for global temperature change scenarios sorted by size. The glaciers are binned according to their initial glacier area and the numbers below each bin (shown in gray) refer to the percentage of the total number of glaciers in 2015 in each bin.

Regional mass changes

Regional variations exist in the glacier mass change projections (Fig. 3). Alaska is the largest regional contributor to global mean sea level rise from 2015 to 2100 (fig. S2), peaking at 0.33 to 0.44 mm SLE yr^{-1} between 2030 and 2060 depending on the temperature change scenario, before decreasing to 0.13 to 0.28 mm SLE yr^{-1} by 2100 (fig. S1). Greenland Periphery, Antarctic and Subantarctic, Arctic Canada North, and Arctic Canada South contribute 12, 10, 10, and 9% to projected sea level rise, respectively.

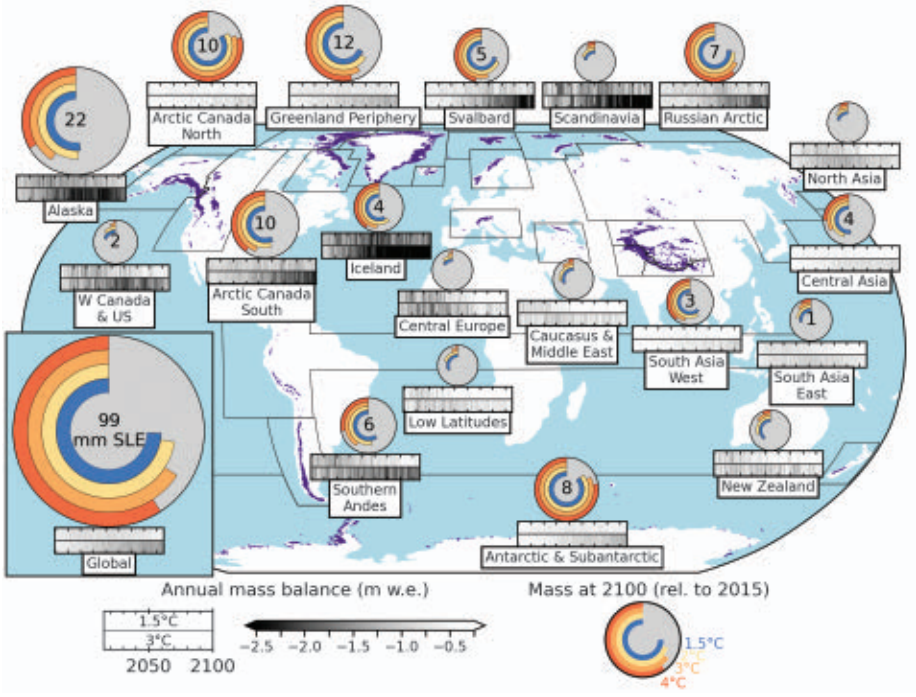


Fig. 3. Regional glacier mass change and contributions to sea level rise from 2015 to 2100. Discs show global and regional projections of glacier mass remaining by 2100 relative to 2015 for global mean temperature change scenarios. Discs are scaled based on each region's contribution to global mean sea level rise from 2015 to 2100 for the +2°C scenario by 2100 relative to preindustrial levels, and nested rings are colored by temperature change scenarios showing normalized mass remaining in 2100. Regional sea level rise contributions >1 mm SLE for the +2°C scenario are printed in the center of each disc. The horizontal bars below each disc show time series of area-averaged annual mass balance from 2015 to 2100 for +1.5°C (top bar) and +3°C (bottom bar) scenarios. The colorbar is saturated at −2.5 m w.e., but minimum annual values reach −4.2 m w.e. in Scandinavia. Time series of regional relative mass change and regional area-averaged mass change are shown in figs. S3 and S4.

Collectively, these five regions account for 60 to 65% of the total glacier contribution to sea level rise. For Greenland Periphery, Arctic Canada North, and Arctic Canada South, the rate of the contribution to sea level rise is almost insensitive to temperature change below +2°C but steadily increases through 2100 for the other temperature change scenarios. For the +3°C and +4°C scenarios, the rate of sea-level rise from Greenland Periphery, Antarctic and Subantarctic, and Arctic Canada North each nearly equal or exceed Alaska near the end of the century, with Antarctic and Subantarctic and Arctic Canada North accelerating throughout the 21st century. Because projected glacier mass loss includes both the instantaneous response of glaciers to climate forcing and the delayed response based on the extent of disequilibrium to longer-term climatic conditions (*20*), these regions with large glaciers will continue losing mass beyond 2100, especially for higher temperature change scenarios.

Western Canada and US, South Asia East, Scandinavia, North Asia, Central Europe, Low Latitudes, Caucasus and Middle East, and New Zealand are projected to lose 60 to 100% of their glacier mass depending on the temperature change scenario (Fig. 3 and fig. S3). The temperature change scenario thus has a major impact on the mass loss, in some cases determining whether the complete deglaciation of regions occurs by the end of the 21st century. Although these regions are not significant contributors to sea level rise, people in these regions will need to adapt to changes in seasonal and annual runoff as the additional water provided by glacier net mass loss will decline before 2050 as the glaciers retreat (figs. S5 and S8). In High Mountain Asia, the timing of maximum rates of mass loss varies, with South Asia East peaking between 2025 and 2030, Central Asia between 2035 and 2055, and South Asia West between 2050 and 2075, depending on the temperature change scenario.

Regional sensitivity to temperature change

The sensitivity of the glacierized regions to changes in global mean temperature depends on the region's current glacier mass and mass change rates; regional temperature anomalies relative to the global mean (Fig. 4), such as those associated with Arctic amplification (*21*); the climatic setting (maritime versus continental); sensitivity to precipitation falling as rain instead of snow; and elevation feedbacks due to different types of glaciers (e.g., ice caps versus valley glaciers) (*22*). Projected mass loss is linearly related to global mean temperature increase, especially for larger glacierized regions, consistent with a recent study (*6*). This strong relationship highlights that every fraction of a degree of temperature increase substantially affects glacier mass loss. The smallest glacierized regions by mass, including Central Europe, Scandinavia, Caucasus and Middle East, North Asia, Western Canada and US, Low Latitudes, and New Zealand, will experience near-complete deglaciation around +3°C. These regions are thus highly sensitive to global mean temperature increases between 1.5 and 3°C and have a nonlinear response above 3°C of warming.

The strength of the linear relationship varies among regions, which reflects differences in the regional temperature anomalies from the ensemble of GCMs (evident from the larger standard deviations given in Fig. 4 and fig. S9). Regions like Alaska, Southern Andes, and Central Asia have less scatter, indicating less variation in the regional temperature anomaly and thereby a more consistent response to climate forcing (mean R^2 = 0.78). Other regions like the Russian Arctic, Svalbard, and Iceland have more variation in the regional temperature anomaly and thus a weaker linear relationship (mean R^2 = 0.50) as well as considerable variation in projected precipitation (fig. S10). Future work using regional climate projections may better resolve high-mountain climatic conditions and refine projections in these regions.

Spatially resolved projections at the glacier scale

Our projections reveal notable spatial variation in glacier mass loss at the local scale for the temperature change scenarios (Fig. 5). All regions are projected to lose some glaciers completely, primarily smaller ice masses, with the higher temperature change scenarios revealing significantly more mass loss and the deglaciation of greater areas (figs. S11 to S13). Although Central Europe, Caucasus and Middle East, North Asia, and Western Canada and US are projected to experience widespread deglaciation for the +2°C scenario, our results also reveal where remaining glaciers will be concentrated at the end of this century. Besides the Karakoram and Kunlun in High Mountain Asia, the remaining mass is primarily located in southeastern Alaska, Arctic Canada North, Svalbard, the Russian Arctic, Greenland Periphery, and Antarctic and Subantarctic. Given that these regions constitute a significant number of marine-terminating glaciers,

accounting for frontal ablation is critical over the next century and beyond.

Importance of marine-terminating glaciers

Marine-terminating glaciers represent 40% of the total present-day global glacier area (*23*), and this percentage reaches 99% for the Antarctic and Subantarctic region. Most previous global glacier projections do not explicitly account for frontal ablation (*2*) and instead implicitly account for it by increasing melt rates, thereby poorly accounting for dynamical feedbacks associated with the glacier's evolution. Our model couples a frontal ablation parameterization with a flowline model and uses a state-of-the-art calibration scheme, ice thickness inversion method, and geodetic mass balance and frontal ablation calibration data (see Methods). These features enable us to project changes of individual marine-terminating glaciers and determine if and when they become land-terminating (fig. S14). Separate simulations including and excluding frontal ablation, with model parameters calibrated separately for both, are used to quantify the impact of accounting for frontal ablation on projections.

Counterintuitively, we estimate that accounting for frontal ablation reduces the glacier contribution to mean sea level rise from 2015 to 2100 by 2% for each temperature change scenario, compared with models not including frontal ablation. From 2015 to 2100 frontal ablation accounts for 91 ± 10 Gigatons (Gt) yr^{-1} (+1.5°C) to 88 ± 8 Gt yr^{-1} (+4°C) of the total glacier mass loss globally (figs. S15 to S18). For the +2°C scenario, the rate of mass loss due to frontal ablation diminishes over the century from 115 ± 11 Gt yr^{-1} in 2000 to 2020 to 75 ± 8 Gt yr^{-1} in 2080 to 2100. Diminished mass losses from frontal ablation of marine-terminating glaciers reflect their thinning, retreat onto land (44 to 57% of all marine-terminating glaciers) (fig. S19), and reduced ice flux into the ocean, which occurs for all temperature change scenarios. The relative contribution of frontal ablation to total ablation (i.e., frontal ablation plus melt) ranges from 11% (+1.5°C) to 8% (+4°C) for 2015 to 2100, diminishing for higher temperature change scenarios due to increases in melt. Regionally, the relative contribution of frontal ablation for all temperature change scenarios is greatest in Antarctic and Subantarctic (34%), the Russian Arctic (34%), and Svalbard (17%) (figs. S15 to S18).

The impact of not accounting for frontal ablation on relative mass loss (i.e., glacier mass loss by 2100 relative to 2015) varies greatly by region (fig. S20). For Alaska and Svalbard, excluding frontal ablation increases relative mass loss at 2100 by 2 to 8% depending on the temperature change scenario. The Russian Arctic varies from a 2% reduction (+1.5°C) to a 5% increase (+4°C). Arctic Canada, Greenland Periphery, and Southern Andes see almost no difference (±2%), and Antarctic and Subantarctic see a 0 to 2% decrease in relative mass loss. These results highlight the complex response of marine-terminating glaciers, which are dependent on the frontal ablation rate, glacier geometry, and surface mass balance. In the Antarctic and Subantarctic, we find excluding frontal ablation decreases the regional relative mass loss, as mass loss due to frontal ablation is greater than the increased melt when frontal ablation is excluded. Conversely, in Alaska and Svalbard, the regional relative mass loss increases when frontal ablation is excluded, as mass loss due to frontal ablation is less than the increased melt when frontal ablation is excluded.

Importance of debris-covered glaciers

Debris currently covers 4 to 7% of the global glacier area (*24*, *25*). A thin layer of debris (<3 to 5 cm) enhances surface melt, whereas a thick layer insulates the underlying ice and reduces melt (*26*). The spatial distribution of debris thickness can cause debris-covered glaciers to

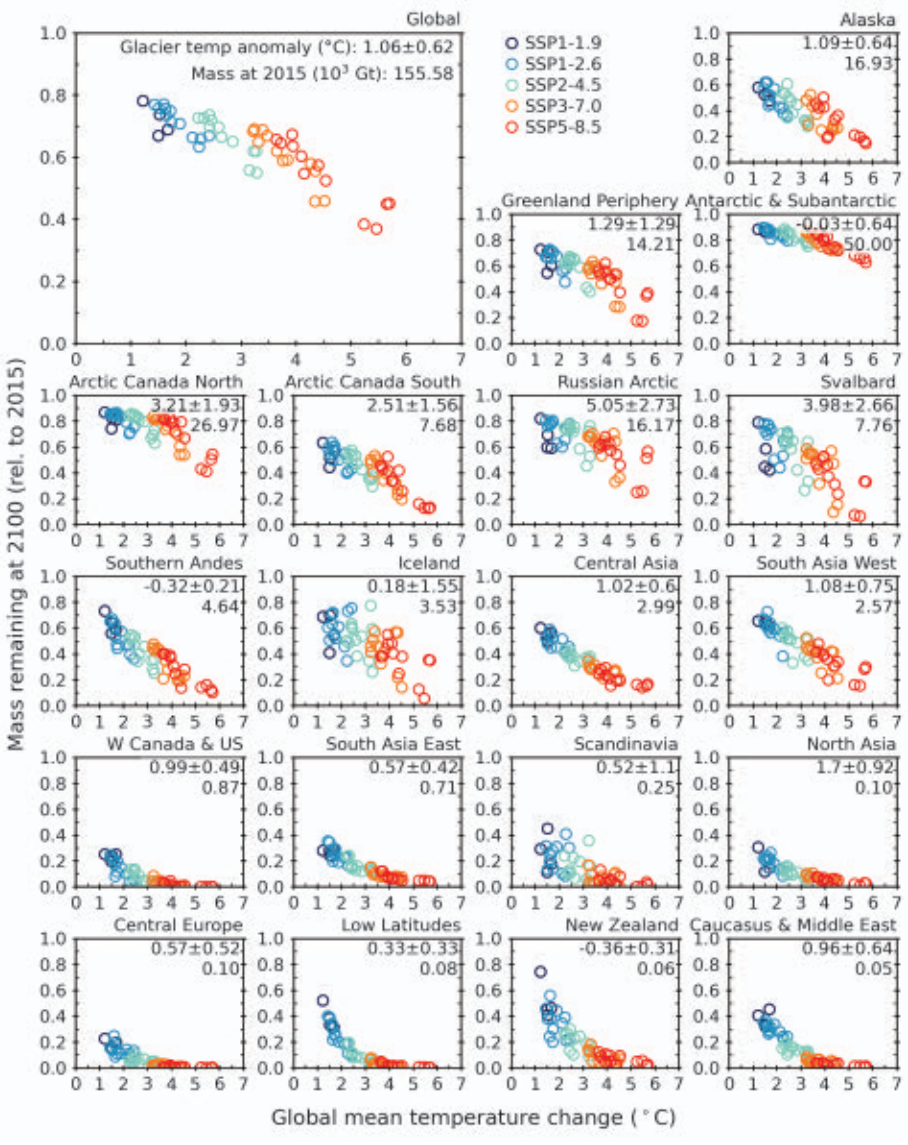


Fig. 4. Fraction of global and regional mass remaining at 2100, relative to 2015, as a function of global mean temperature change by 2100 relative to preindustrial levels. Each marker represents results from one GCM and SSP. Numbers indicate median temperature anomalies (± standard deviation) (°C) over glacierized areas, relative to the mean temperature change over the entire globe at 2100 relative to preindustrial levels, for all GCMs and scenarios, and the glacier mass at 2015 (10^3 Gt). Negative values indicate that some regions warm less than the global average. Regions are ordered by their total mass loss.

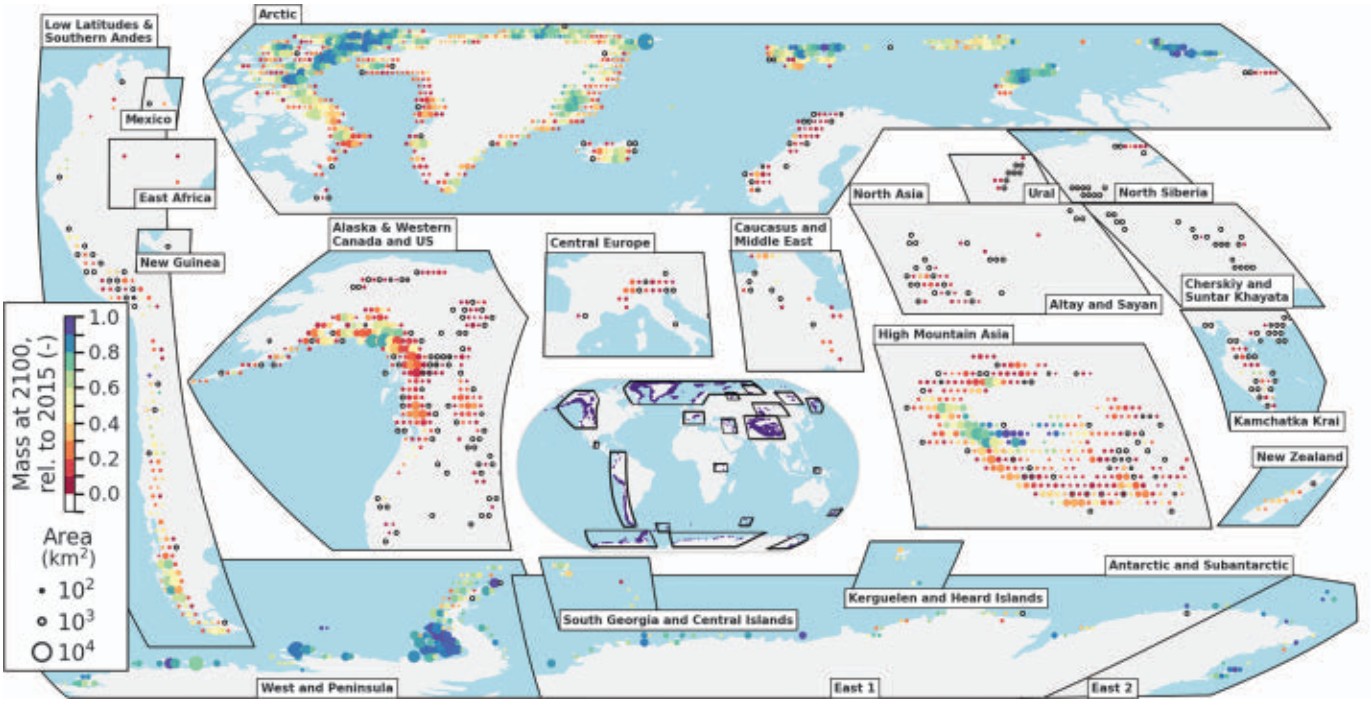


Fig. 5. Spatial distribution of glacier mass remaining by 2100 for the +2°C scenario. The ensemble median glacier mass remaining by 2100 (relative to 2015) for the +2°C (above preindustrial levels) global mean temperature change scenario. Tiles are aggregated by 1° by 1° below 60° latitude, 2° by 1° between 60° and 74° latitude and 2° by 2° above 74° latitude to represent ~10,000 km^2 each. Circles are scaled based on simulated glacierized area in 2015 and are colored by normalized mass remaining. Regions that have experienced complete deglaciation by 2100 are shown in white and outlined in black. High Mountain Asia refers to Central Asia, South Asia West, and South Asia East. Specific subregions are noted by labels on the bottom of inset figures. Additional temperature change scenarios (+1.5°C, +3°C, and +4°C) are shown in figs. S11 to S13.

Table 1. Projected global glacier mass loss and glacier contribution to sea level rise. Results are shown for RCP and SSP scenarios at 2100, relative to 2015, from this study and recent multimodel studies (*2*, *6*). "Uncorrected" refers to projections that assume mass losses below sea level contribute to sea level rise, consistent with assumptions in recent multimodel studies. Note that uncertainty associated with the multimodel studies is expressed as 90% CI, whereas this study reports ensemble median and 95% CI. Regional comparisons are shown in tables S1 and S2.

Global glacier contribution to sea level rise from 2015 to 2100 (mm SLE)

Study	RCP2.6	RCP4.5	RCP8.5	SSP1-2.6	SSP2-4.5	SSP5-8.5
This study	90 ± 36	114 ± 44	163 ± 53	98 ± 38	116 ± 51	166 ± 83
This study (uncorrected)	106 ± 37	132 ± 47	187 ± 61	115 ± 42	135 ± 57	192 ± 97
Marzeion *et al*. (*2*)	79 ± 57	119 ± 66	159 ± 86	-	-	-
Edwards *et al*. (*6*)	-	-	-	80 ± 35	119 ± 39	159 ± 47

Global glacier mass loss, relative to 2015 (%)

Study	RCP2.6	RCP4.5	RCP8.5	SSP1-2.6	SSP2-4.5	SSP5-8.5
This study	26 ± 8	31 ± 10	43 ± 13	28 ± 9	32 ± 12	44 ± 20
Marzeion *et al*. (*2*)	18 ± 13	27 ± 15	36 ± 20	-	-	-

develop stagnant glacier tongues and eventually separate from the active part of the glacier (*27*, *28*). Our representation of debris and glacier dynamics enables us to simulate these complex feedbacks, including reduced melt at glacier termini where debris is thick (fig. S21). We thus produce a set of global glacier projections that account for debris and compare these to separate simulations that exclude debris (i.e., treating the debris as clean ice) to quantify the insulating effect that debris has on glacier projections.

The impact of debris on relative mass loss varies greatly spatially and temporally (fig. S22) with the most significant differences occurring around mid-century in New Zealand and South Asia East. In these regions, the insulating effect of debris reduces net mass loss by 9 to 13% depending on the temperature change scenario, although the differences are less than 5% by 2100. Alaska, the largest region by mass with considerable debris cover (>5% by area), sees a reduction of 5% around 2060 and 3% by 2100. Other regions with considerable debris cover (>5% by area), including Western Canada and US, Central Europe, Caucasus and Middle East, and Low Latitudes, see a reduction in mass loss of less than 5% around mid-century and no difference (±1%) by 2100. The inclusion of debris thus delays mass loss over the century

especially at local scales but has little impact on sea level rise and the number of glaciers lost by 2100. The limited impact in most regions shows that the insulating effect of debris is unable to offset the increased melt for the various temperature change scenarios.

Comparison with previous projections

For comparison with recent multimodel studies (*2*, *6*), we also report our projections for the RCPs and SSPs. Our global projections of glacier contribution to sea level rise for 2015 to 2100 range from 90 ± 36 mm SLE (RCP2.6) to 163 ± 53 mm SLE (RCP8.5) and 98 ± 38 mm SLE (SSP1-2.6) to 166 ± 83 mm SLE (SSP5-8.5), respectively (Table 1). These projections include a correction (reduction) of 17 to 24 mm SLE, which accounts for the mass loss of ice from marine-terminating glaciers that is below sea level and therefore will not contribute to global mean sea level rise—an important difference compared with current multimodel studies (*2*, *6*), which do not account for this. Even with this correction, for the low emissions scenarios our RCP2.6 projections are 11 mm SLE (14%) greater than that of Marzeion *et al.* (*2*), and our SSP1-2.6 projections are 18 mm SLE (23%) greater than that of Edwards *et al.* (*6*). For the mid-range (RCP4.5 and SSP2-4.5) and high (RCP8.5, SSP5-8.5) emissions scenarios, our projections are within ±7 mm SLE of both studies.

Not correcting for the loss of ice below sea level, our projections of glacier contribution to sea level rise from 2015 to 2100 are 11 to 44% greater than these multimodel estimates (*2*, *6*) for all emission scenarios. We attribute these differences to the global mass balance data we used for calibration, which include an accelerated trend in mass loss from 2000 to 2020 (*1*), as well as the improved representation of physical processes in our model.

Globally, we predict glaciers will lose 26 ± 8% (RCP2.6) to 43 ± 13% (RCP8.5) and 28 ± 9% (SSP1-2.6) to 44 ± 20% (SSP5-8.5) of their mass by 2100, relative to 2015. Our projected relative mass losses are 4 to 8% greater than current multimodel estimates (*2*). Regionally, the most significant differences occur in Alaska, Arctic Canada South, South Asia East, and Southern Andes, where we predict 11 to 23% more relative mass loss (table S1). In Alaska, we estimate 22% (RCP2.6) to 23% (RCP8.5) more relative mass loss compared with the multimodel estimates (*2*) and find a peak in the net mass loss rate in the middle of the century, in contrast to the peak net mass loss rate at the end of the century from the multimodel estimates (*2*).

A comparison of our projections from the ensembles of RCPs and SSPs used in this study reveals that glacier contribution to sea level rise is 2 to 9% greater for SSPs than the corresponding RCPs. These differences are a result of the SSPs simulating greater temperature increases for the same radiative forcing as the RCPs (*29*, *30*). Our ensembles reflect this higher warming sensitivity as the SSPs are on average 0.14 to 0.25°C warmer than their corresponding RCPs. Considering the high sensitivity of global and regional glacier mass loss to small temperature increases revealed by our study, the higher warming sensitivity of the SSPs will substantially affect the projected glacier contribution to sea level rise as well as the number of glaciers anticipated to be lost.

Summary and way forward

Our projections reveal a strong linear relationship between global mean temperature increase and glacier mass loss, with the smallest glacierized regions having a nonlinear relationship beyond +3°C as they experience near-complete deglaciation. This strong relationship at global and regional scales highlights that every increase in temperature has significant consequences with respect to glacier contribution to sea level rise, the loss of glaciers around the world, and changes to hydrology, ecology, and natural hazards. Regardless of the temperature change scenario, all regions will experience considerable deglaciation at local scales with roughly half of the world's glaciers by number projected to be lost by 2100, even if temperature increase is limited to +1.5°C. Based on the most recent climate pledges from COP26, global mean temperature is estimated to increase by +2.7°C (*31*), which would result in much greater glacier contribution to sea level rise (115 ± 40 mm SLE) and the near-complete deglaciation of entire regions including Central Europe, Western Canada and US, and New Zealand (Fig. 5 and figs. S11 to S13) compared with the Paris Agreement. The rapidly increasing glacier mass losses as global temperature increases beyond +1.5°C stresses the urgency of establishing more ambitious climate pledges to preserve the glaciers in these mountainous regions.

REFERENCES AND NOTES

1. R. Hugonnet *et al.*, *Nature* **592**, 726–731 (2021).
2. B. Marzeion *et al.*, *Earths Futur.* **8**, (2020).
3. W. W. Immerzeel *et al.*, *Nature* **577**, 364–369 (2020).
4. M. Huss, R. Hock, *Nat. Clim. Chang.* **8**, 135–140 (2018).
5. S. Harrison *et al.*, *Cryosphere* **12**, 1195–1209 (2018).
6. T. L. Edwards *et al.*, *Nature* **593**, 74–82 (2021).
7. F. Maussion *et al.*, *Geosci. Model Dev.* **12**, 909–931 (2019).
8. M. Huss, R. Hock, *Front. Earth Sci.* **3**, (2015).
9. R. Hock *et al.*, *J. Glaciol.* **65**, 453–467 (2019).
10. B. Osmanoğlu, M. Braun, R. Hock, F. J. Navarro, *Ann. Glaciol.* **54**, 111–119 (2013).
11. B. Osmanoglu, F. J. Navarro, R. Hock, M. Braun, M. I. Corcuera, *Cryosphere* **8**, 1807–1823 (2014).
12. M. Minowa, M. Schaefer, S. Sugiyama, D. Sakakibara, P. Skvarca, *Earth Planet. Sci. Lett.* **561**, 116811 (2021).
13. W. Kochtitzky *et al.*, *Nat. Commun.* **13**, 5835 (2022).
14. D. R. Rounce *et al.*, *J. Glaciol.* **66**, 175–187 (2020).
15. D. R. Rounce, R. Hock, D. E. Shean, *Front. Earth Sci.* **7**, 331 (2020).
16. K. Hutter, *J. Glaciol.* **27**, 39–56 (1981).
17. D. R. Rounce *et al.*, *Geophys. Res. Lett.* **48**, GL091311 (2021).
18. D. Farinotti *et al.*, *Nat. Geosci.* **12**, 168–173 (2019).
19. R. Hock *et al.*, in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner *et al.*, Eds. (Cambridge University Press, 2019).
20. H. Zekollari, M. Huss, D. Farinotti, *Cryosphere* **13**, 1125–1146 (2019).
21. F. Pithan, T. Mauritsen, *Nat. Geosci.* **7**, 181–184 (2014).
22. J. Bolibar, A. Rabatel, I. Gouttevin, H. Zekollari, C. Galiez, *Nat. Commun.* **13**, 409 (2022).
23. RGI Consortium, Randolph glacier inventory - A dataset of global glacier outlines, Version 6.0, GLIMS (2017); https://doi.org/10.7265/4m1f-gd79
24. D. Scherler, H. Wulf, N. Gorelick, *Geophys. Res. Lett.* **45**, 11798–11805 (2018).
25. S. Herreid, F. Pellicciotti, *Nat. Geosci.* **13**, 621–627 (2020).
26. G. Østrem, *Geogr. Ann.* **41**, 228–230 (1959).
27. D. I. Benn *et al.*, *Earth Sci. Rev.* **114**, 156–174 (2012).
28. A.V. Rowan *et al.*, *J. Geophys. Res. Earth Surf.* **126**, (2021).
29. K. B. Tokarska *et al.*, *Sci. Adv.* **6**, aaz9549 (2020).
30. K. Wyser, E. Kjellström, T. Koenigk, H. Martins, R. Döscher, *Environ. Res. Lett.* **15**, 054020 (2020).
31. UNEP, "Emissions Gap Report 2021" (UNEP, 2021); https://www.unep.org/emissions-gap-report-2021).

ACKNOWLEDGMENTS

This work was supported in part by the high-performance computing and data storage resources operated by the Research Computing Systems Group at the University of Alaska Fairbanks Geophysical Institute. This text reflects only the author's view and funding agencies are not responsible for any use that may be made of the information it contains. **Funding:** This work was funded by the following: National Aeronautics and Space Administration grant 80NSSC20K1296 (to D.R. and Re.Ho.); National Aeronautics and Space Administration grant 80NSSC20K1595 (to D.R. and Re.Ho.); National Aeronautics and Space Administration grant 80NSSC17K0566 (to D.R. and Re.Ho.); National Aeronautics and Space Administration grant NNX17AB27G (to D.R. and Re.Ho.); Norwegian Research Council project #324131 (to Re.Ho.); Tula Foundation and Canada Research Chairs (to B.M.); National Sciences and Engineering Research Council of Canada (to B.M. and Lu.Co.); Vanier Graduate Scholarship (to W.K.); Swiss National Science Foundation project 184634 (to Ro.Hu., M.H., Lo.Co., and D.F.); ArcticNet Network of Centres of Excellence Canada (to Lu.Co.); University of Ottawa, University Research Chair program (to Lu.Co.); European Union's Horizon 2020 research and innovation programme grant 101003687 (to F.M.); Austrian Science Fund (FWF) grant P30256 (to F.M.); French Space Agency CNES (to E.B. and Ro.Hu.) **Author contributions:** Conceptualization: D.R. and Re.Ho. Data curation: D.R. Formal analysis: D.R. Funding acquisition: D.R., Re.Ho., M.H., D.F., E.B., B.M., and Lu.Co. Investigation: D.R. Methodology: D.R., F.M., and Re.Ho. Project administration: D.R. and Re.Ho. Resources: D.R., F.M. (glacier data); Ro.Hu., M.H., E.B., D.F., B.M., and R.M. (mass balance data); Lo.Co. (climate data); W.K. and Lu.Co. (frontal ablation data); Software: D.R. (PyGEM); F.M. (OGGM); D.B. (emulators); Visualization: D.R., Re.Ho, and Ro.Hu. Writing – original draft: D.R. Writing – review and editing: all authors, especially Re.Ho. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** The datasets generated for this study can be found in the National Snow and Ice Data Center (NSIDC) at https://nsidc.org/data/hma2_ggp/versions/1. The model code is publicly available at https://github.com/drounce/PyGEM and https://github.com/OGGM/oggm. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo1324
Methods
References (*32*–*41*)
Figs. S1 to S28
Tables S1 to S5

Submitted 15 January 2022; accepted 14 November 2022
10.1126/science.abo1324

BRAIN ANATOMY

A mesothelium divides the subarachnoid space into functional compartments

Kjeld Møllgård[1]*†, Felix R. M. Beinlich[2]†, Peter Kusk[2]†, Leo M. Miyakoshi[2]†, Christine Delle[2], Virginia Plá[2], Natalie L. Hauglund[2], Tina Esmail[2], Martin K. Rasmussen[2], Ryszard S. Gomolka[2], Yuki Mori[2], Maiken Nedergaard[3]*

The central nervous system is lined by meninges, classically known as dura, arachnoid, and pia mater. We show the existence of a fourth meningeal layer that compartmentalizes the subarachnoid space in the mouse and human brain, designated the subarachnoid lymphatic-like membrane (SLYM). SLYM is morpho- and immunophenotypically similar to the mesothelial membrane lining of peripheral organs and body cavities, and it encases blood vessels and harbors immune cells. Functionally, the close apposition of SLYM with the endothelial lining of the meningeal venous sinus permits direct exchange of small solutes between cerebrospinal fluid and venous blood, thus representing the mouse equivalent of the arachnoid granulations. The functional characterization of SLYM provides fundamental insights into brain immune barriers and fluid transport.

Emerging evidence supports the concept that cerebrospinal fluid (CSF) acts as a quasi-lymphatic system in the central nervous system (*1*). Cardiovascular pulsatility drives CSF inflow along periarterial spaces into deep brain regions (*2*, *3*), where CSF exchange with interstitial fluid, facilitated by glial aquaporin 4 (AQP4) water channels (*4*), takes place. Fluid and solutes from the neuropil are cleared along multiple routes, including perivenous spaces and cranial nerves, for ultimate export to the venous circulation via meningeal and cervical lymphatic vessels (*5*, *6*). CSF reabsorption may also occur at the sinuses via the arachnoid granulations—although this has not been described in rodents (*7*–*10*). Despite the efforts dedicated to studying CSF flow along the glymphatic-lymphatic path, it remains to be determined how CSF is transported within the large cavity of the subarachnoid space (*11*, *12*). In this study, we explored how CSF and immune cell trafficking are organized within the subarachnoid space surrounding the brains of mice and humans.

The meningeal membranes were first analyzed by in vivo two-photon microscopy in the somatosensory cortex of Prox1-EGFP$^+$ reporter mice (Prox1, prospero homeobox protein 1; EGFP, enhanced green fluorescent protein). Prox1 is a transcription factor that determines lymphatic fate (*13*, *14*). Second harmonic generation was used to visualize unlabeled collagen fibers, while the vascular volume was labeled by a Cascade Blue conjugated dextran, and astrocytes were labeled by sulforhodamine 101 (SR101, intraperitoneally) (*15*, *16*). Below the parallel-oriented collagen bundles in dura, we noted a continuous monolayer of flattened Prox1-EGFP$^+$ cells intermixed with loosely organized collagen fibers. This subarachnoid lymphatic-like membrane (SLYM) divides the subarachnoid space into an outer superficial compartment and an inner deep compartment lining the brain (Fig. 1A). Quantitative in vivo analysis of the somatosensory cortex revealed that the thickness of SLYM itself was 14.2 ± 0.5 μm, hence thinner than dura (21.8 ± 1.3 μm, n = 6 mice). The dura vasculature is surrounded by collagen fibers, whereas SLYM covers the subarachnoid vessels. The organization and calibers of the two sets of vasculature also exhibit distinct differences (Fig. 1, B and C).

A key question is whether SLYM constitutes an impermeable membrane that functionally compartmentalizes the subarachnoid space. To test this, Prox1-EGFP$^+$ mice were first injected with 1-μm microspheres conjugated to a red fluorophore into the subdural outer superficial compartment of the subarachnoid space along with an injection of 1-μm microspheres conjugated to a blue fluorophore distributed within the inner deep subarachnoid space compartment by cisterna magna injection (Fig. 2A). In vivo two-photon microscopy showed that the red microspheres were confined to the outer superficial compartment, whereas the blue microspheres remained trapped in the inner deep subarachnoid space compartment. Quantitative analysis showed that the 1-μm microspheres did not cross SLYM from either side. Yet, many solutes in CSF, such as cytokines and growth factors, are considerably smaller than 1 μm in diameter (*17*). Therefore, we sought to determine whether a small tracer could pass through SLYM. In these experiments, tetramethylrhodamine (TMR)–dextran (3 kDa) was administered into the deep inner subarachnoid space via the cisterna magna in Prox1-EGFP$^+$ mice. In six mice, the small tracer did not cross the EGFP-expressing

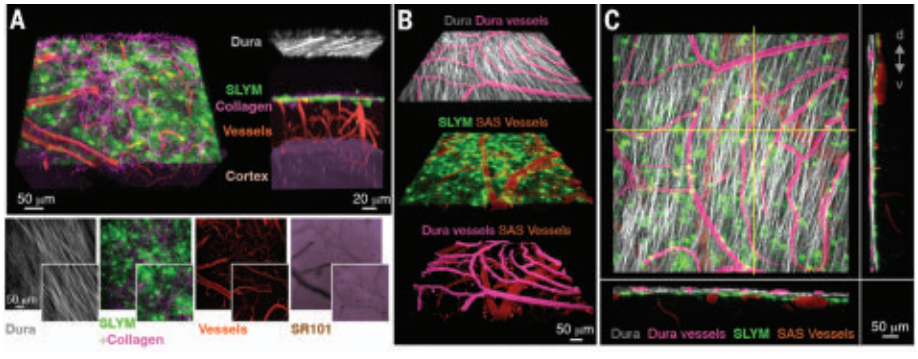


Fig. 1. In vivo imaging depicts a fourth meningeal layer. (**A**) In vivo two-photon imaging of Prox1-EGFP$^+$ reporter mice viewed through a closed cranial window placed over the somatosensory cortex. Maximum projection and three-dimensional (3D) views depict the spatial distribution of dura mater collagen fibers (gray) detected by second harmonic generation. Prox1-EGFP$^+$ cells (green) intermixed with the irregular sparse collagen fibers (purple) localized below dura. This subarachnoid lymphatic-like membrane is abbreviated SLYM. Blood vessels outlined by Cascade Blue conjugated dextran (red, 10 kDa, iv) are located at the cortical surface. (Inset) A lateral view of the 3D reconstruction with all the layers displayed individually along the z axis to facilitate spatial comprehension. (**B**) Two-photon imaging over the sensorimotor cortex in a Prox1-EGFP reporter mouse. The vasculature was outlined by intravenous injection of TMR-dextran (2000 kDa), and z-stacks were collected. Representative 3D reconstruction of the z-stacks. The vasculature in dura (magenta) is embedded in collagen fibers (white). In contrast, the vasculature in the subarachnoid space (red) is overlaid by SLYM (green). (**C**) Orthogonal sections through the z-stack show that the vasculature in dura is surrounded by collagen fibers. SLYM is located beneath dura, in close apposition with the large-caliber subarachnoid vessels.

[1]Department of Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. [2]Division of Glial Disease and Therapeutics, Center for Translational Neuromedicine, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. [3]Division of Glial Disease and Therapeutics, Center for Translational Neuromedicine, University of Rochester Medical Center, Rochester, NY 14642, USA.
***Corresponding author. Email: nedergaard@urmc.rochester.edu (M.N.); kjm@sund.ku.dk (K.M.)**
†These authors contributed equally to this work.

SLYM (Fig. 2B and fig. S1). Yet, in mice with dural damage and leakage of CSF, the tracer was observed on both sides of the EGFP$^+$ membrane (fig. S1). Thus, SLYM divides the subarachnoid space into an upper superficial and a lower deep compartment for solutes ≥3 kDa. SLYM is therefore a barrier that limits the exchange of most peptides and proteins, such as amyloid-β and tau, between the upper and lower subarachnoid space compartments.

Live brain imaging avoids fixation artifacts (*18*) but cannot immunophenotypically characterize the meningeal membranes. To preserve the integrity of the meningeal membranes, sections were next obtained from whole heads of Prox1-EGFP$^+$ mice. Immunohistochemistry revealed that Prox1-EGFP$^+$ cells lined the ventral parts of the entire brain surface (Fig. 3A). Immunolabeling showed that the Prox1-EGFP$^+$ SLYM cells were positive for another lymphatic marker, podoplanin (PDPN) (*19*), but not for the lymphatic vessel endothelial receptor 1 (LYVE1) (*20*) (Fig. 3, A, lower right panels, and D). SLYM also labeled for the cellular retinoic acid–binding protein 2 (CRABP2) (Fig. 3, A and D), which is restrictively expressed in dural and arachnoid cells during early development (*21*). In contrast to SLYM, lymphatic vessels in dura were positive for all the classical lymphatic antigens, Prox1-EGFP$^+$, PDPN$^+$, LYVE1$^+$, and VEGFR3$^+$, but was CRABP2$^-$ (fig. S2). Notably, analysis of adult human cerebral cortex depicted that above the pia mater, a CRABP2$^+$/PDPN$^+$ membrane was present in the entire subarachnoid space (Fig. 3, B and C). Thus, SLYM also surrounds the human brain. We

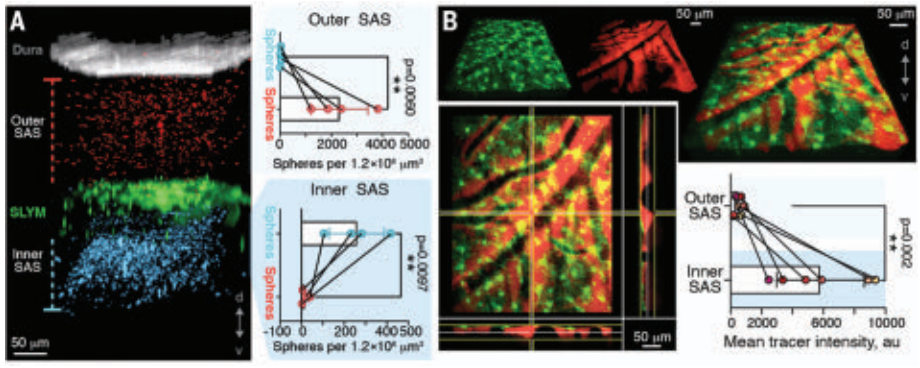


Fig. 2. SLYM represents a barrier that subdivides the subarachnoid space into two compartments. (**A**) Representative image of a 3D view of maximum projection collected after dual injections of red microspheres (red, 1 µm) into the outer superficial subarachnoid space (subdural) and blue microspheres delivered into the inner deep subarachnoid space by cisterna magna injection (blue, 1 µm) in a Prox1-EGFP$^+$ mouse. Graphs show a comparison of red microspheres versus blue microspheres detected in both the outer and inner subarachnoid space (SAS). Two-tailed unpaired *t* test; outer SAS, $P < 0.01$; inner SAS, $P < 0.01$; $n = 4$ mice. (**B**) Representative in vivo *z*-stack of a Prox1-EGFP mouse injected with a 3-kDa TMR-conjugated dextran CSF tracer delivered via the cisterna magna. Upper panels depict SLYM (green) and the perivascular distribution of the dextran (red) as well as the two channels merged. Lower panel displays the merge of the two channels and orthogonal optical sections showing that tracer is confined to below the membrane. Graph shows the mean tracer intensity detected below and above the membrane. Two-tailed unpaired *t* test with Welch's correction, $P < 0.01$, $n = 6$ mice. Significance shown as $^{**}P < 0.01$. au, arbitrary units; CM, cisterna magna; d, dorsal; v, ventral.

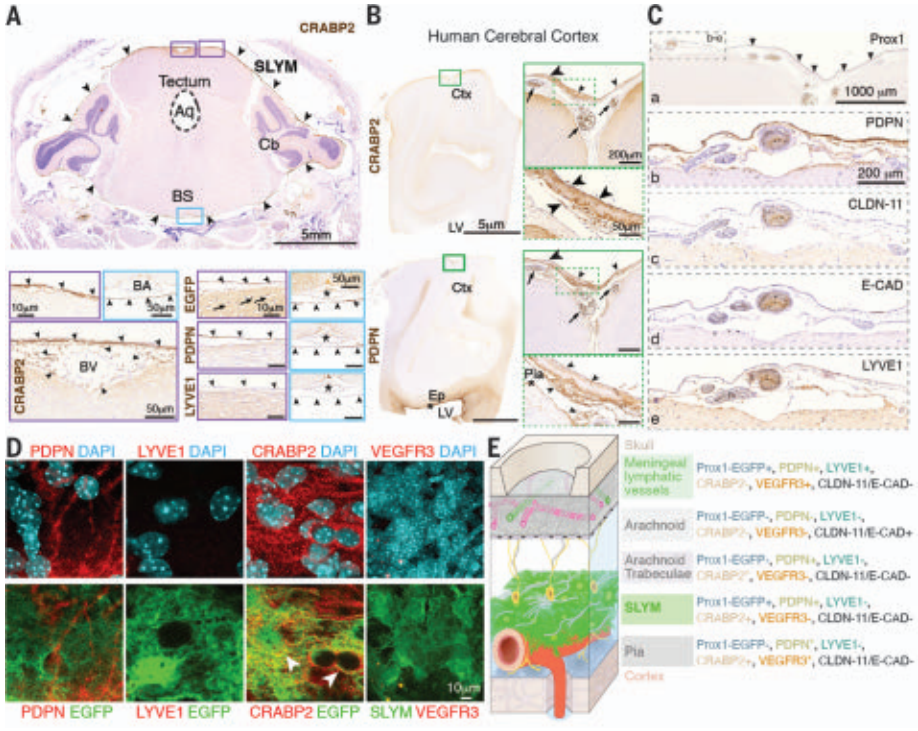


Fig. 3. Immunophenotypic characterization of SLYM in the mouse and human brain. (**A**) Sections of Prox1-EGFP$^+$ mouse brain after decalcification of the whole head counterstained with Mayer's hematoxylin (M-HE, purple) show that SLYM (arrowheads) is positively immunolabeled for CRABP2 (brown) and Prox1-EGFP$^+$/PDPN$^+$/LYVE1$^-$/VEGFR3$^-$ and encases the entire brain, covering its dorsal and ventral portions (purple and blue insets, respectively). (**B**) Adult human brain sections immunolabeled for CRABP2 and PDPN reveal the presence of SLYM (arrowheads) that enwraps the subarachnoid space blood vessels (arrow). Ependymal and pia mater cells are also PDPN$^+$ (asterisks). (**C**) Serial sections of the same adult human material immunolabeled for Prox1, PDPN, CLDN11, E-CAD, and LYVE1. SLYM is indicated by arrowheads. (**D**) Confocal images of SLYM immunolabeling showing positive labeling for PDPN and CRABP2 (both in red). No signal was detected for LYVE1 or VEGFR3. (**E**) Schematic representations of the immunophenotypical characterization of the meningeal layers, meningeal lymphatic vessels, and arachnoid trabeculations. For arachnoid trabeculae, CRABP2* signifies that the trabeculae are CRABP2$^-$ in the outer SAS but CRABP2$^+$ in the inner SAS. For pia, PDPN* indicates that pia is PDPN$^+$ in many regions of pia, but not all. VEGFR3* signifies that pia was VEGFR3$^+$ only in a few regions. Aq, aqueduct; BA, basilar artery; BS, brain stem; BV, blood vessel; Cb, cerebellum; Ctx, cerebral cortex; Ep, ependyma; LV, lateral ventricle.

infer that the SLYM monolayer of Prox1-EGFP$^+$ cells organizes into a membrane rather than vessel structures and exhibits a distinctive set of lymphatic markers (Fig. 3E). To distinguish SLYM from the structures forming the arachnoid mater, we used immunolabeling against claudin-11 (CLDN-11), a main constituent of the tight junctions that create the arachnoid barrier cell layer (ABCL) (*22*). CLDN-11 was densely expressed in ABCL as well as in the stromal cells of the choroid plexus, but SLYM was CLDN-11$^-$ (fig. S3, A and B). Additionally, ABCL was distinctively positive for E-cadherin (E-Cad) (Fig. 3C), as previously reported (*23*, *24*). We also compared SLYM to the arachnoid trabeculae (*25*), collagen-enriched structures that span the subarachnoid space, finding that cells surrounding the arachnoid trabeculae are Prox1-EGFP$^-$/LYVE1$^-$ (fig. S3C). Pial cells covering the cortical surface also exhibited an immune-labeling profile that differed from that of SLYM (figs. S3 and S4). We conclude that SLYM constitutes a fourth meningeal layer surrounding the mouse and human brain displaying lymphatic-like features (Prox1-EGFP$^+$, PDPN$^+$, LYVE1$^-$, CRABP2$^+$, VEGFR3$^-$, CLDN-11$^-$, and E-Cad$^-$) and that SLYM is phenotypically distinct from dura, the arachnoid, and pia mater (Fig. 3E). Interestingly, SLYM expressed PDPN, sharing a trait with the mesothelium lining the body cavities (*26*). Accordingly, we observed PDPN$^+$ cells lining the kidney, as well as PDPN$^+$ podocytes in the kidneys of adult C57BL/6J mice (fig. S5A). In a human fetus, a PDPN$^+$ membrane corresponding to pericardium, pleura, and peritoneum encases the developing heart, lungs, and intestinal tract, respectively. PDPN$^+$ lymphatic vessels were also observed in the lungs and intestinal tract (fig. S5, B and C). Thus, SLYM may represent the brain mesothelium and, as such, covers blood vessels in the subarachnoid space (Fig. 1) (*26*). The mesothelium is present where tissues slide against each other and is believed to act as a boundary lubricant to ease movement (*27*). Physiological pulsations induced by the cardiovascular system, respiration, and positional changes of the head are constantly shifting the brain within the cranial cavity. SLYM may, like other mesothelial membranes, reduce friction between the brain and skull during such movements.

Does SLYM have additional functions? The arachnoid villi and granulations are defined as protrusions of the arachnoid membrane into the lateral walls of the sinus veins and are believed to act as passive filters that drain

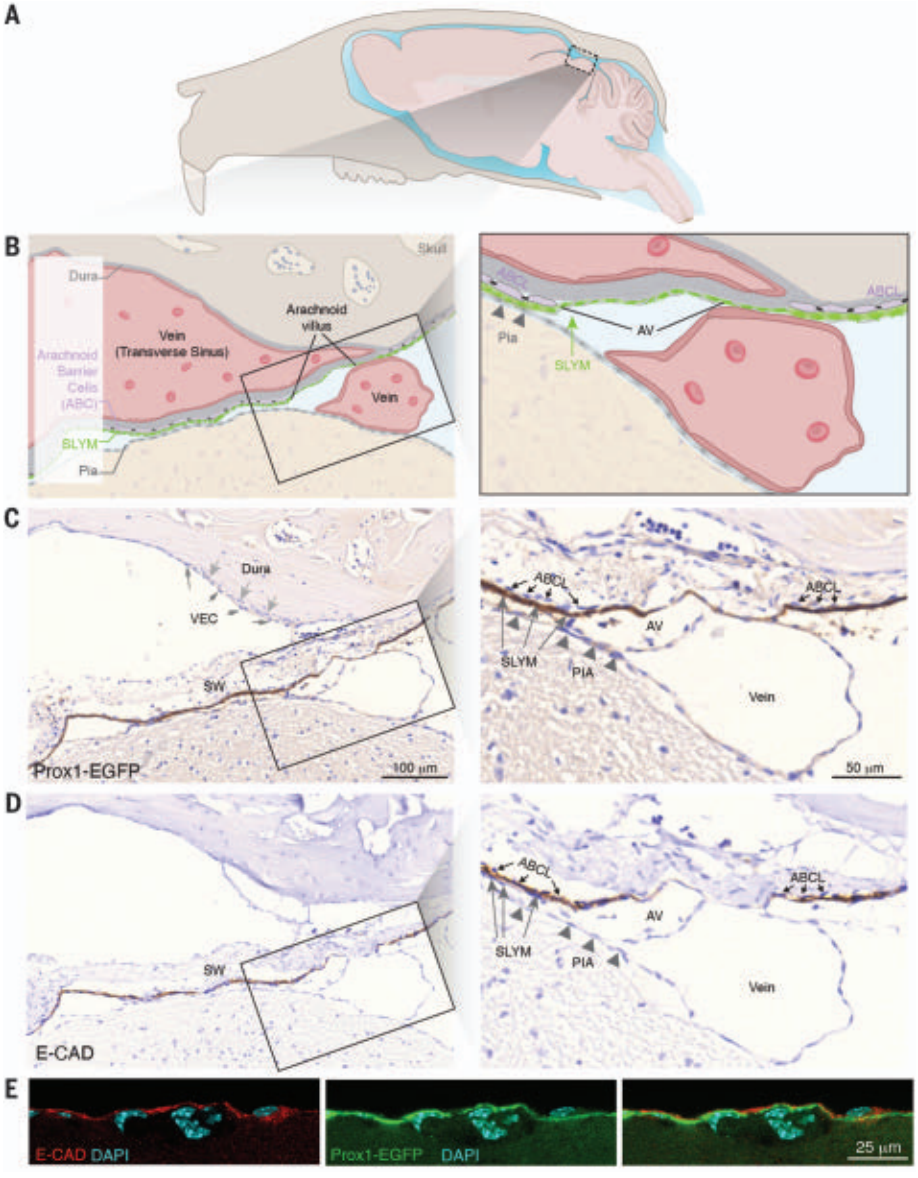


Fig. 4. SLYM forms subarachnoid villus–like structures at the venous sinus walls in mice. (**A** and **B**) Schematic diagrams illustrating the region of interest. (**C** and **D**) Parasagittal consecutive sections from a decalcified mouse whole head stained for (C) the SLYM marker Prox1-EGFP and (D) the arachnoid barrier cell marker E-Cad. Rectangular insets on the left in (B), (C), and (D) are shown in higher magnification on the right. A Prox1-positive arachnoid villus–like structure (AV) and a vein from the dorsal venous system are in direct contact with the transverse sinus wall (SW), which is lacking an intervening ABCL [inset in (C) and (D)]. The ABCL [arrows in inset in (C)] are not stained for Prox1, in contrast to the strongly stained SLYM layer, whereas the opposite pattern of reactivity is depicted in the adjacent section [inset in (D)], where ABCL is positive for E-Cad and SLYM is negative. Arrowheads point to pia. In (C), the narrow dura layer, indicated by small arrows, is facing the venous endothelial layer (VEC), indicated by slender darker arrows. (C) and (D) are the same magnification, as are their insets. (**E**) Confocal imaging of Prox1-EGFP and E-Cad shows that the signals do not colocalize.

CSF from the subarachnoid space into the venous sinus system (*7*–*10*). The arachnoid villi and granulations are present in the brains of humans, primates, and larger animals such as dogs, but not in the brains of rodents (*28*, *29*). We critically reexamined this issue to evaluate the distribution of SLYM in relation to the superior sagittal and transverse sinus. Sections obtained from decalcified heads of Prox1-EGFP$^+$ mice showed that Prox1-EGFP$^+$ SLYM cells often were in direct contact with the venous sinus endothelial cells (Fig. 4A). Thus, the arachnoid barrier cell layer (CLDN-11$^+$/E-Cad$^+$), which normally separates dura from the subarachnoid space, was lacking in discrete areas allowing SLYM to directly contact the venous sinus wall (Fig. 4B). Prox1-EGFP$^+$ SLYM cells were not positive for CLDN-11 or E-Cad, which distinguish the arachnoid barrier cell layer (fig. S6).

Are the close appositions of SLYM and the venous endothelial cells permeable, allowing the exchange of small molecules between blood and CSF? To test this, we used the principles of bioluminescence, wherein the convergence, in the same compartment, of an enzyme with its substrate is needed to trigger light emission. First, we delivered the luciferase enzyme from *Oplophorus gracilirostris* (NanoLuc) fused to the fluorescence tag mNeongreen (GeNL, 44 kDa) (*30*) into CSF via the cisterna magna of wild-type (C57bl/6) mice, and allowed it to circulate for 30 min to ensure thorough distribution by the glymphatic system. The distribution of GeNL was verified by mNeongreen fluorescence. Then, the blood-brain barrier (BBB)–impermeable substrate fluorofurimazine (FFz, 433 Da) (*31*) was administered intravenously (fig. S7, A to C) (*32*). After intravenous injection of FFz, a bright bioluminescence signal catalyzed by GeNL was detected specifically near the large venous sinus wall (fig. S7, A and B). The bioluminescence signal was particularly strong around the confluence of sinuses (fig. S7B). The distribution of the bioluminescence signal was quantified by plotting the mean signal intensity profiles perpendicular to the venous wall of the transverse sinus and superior sagittal sinus. The mean bioluminescence signal profiles intersected with the fluorescence signal profiles of the intravascular tracer (TMR-dextran, 70 kDa) or with shadow imaging of the inverted GeNL signal outlining the vascular wall (fig. S7C). Thus, the bioluminescence signal was restricted to the venous wall of the two major sinuses lacking a BBB (*33*, *34*), consistent with the notion that FFz is BBB-impermeable and requires the catalyzation enzyme NanoLuc to generate photons (fig. S7, A to C). In control experiments, FFz was delivered intravenously, while the GeNL injection into CSF was omitted. In these control experiments, no bioluminescence signal was detected from the exposed cortex, including from the sinus venous wall (fig. S7D). In another set of control experiments, GeNL was injected into the soft ear tissue, while FFz was delivered intravenously. Consistent with the notion that peripheral blood vessels are leaky (*11*), light emission was clearly observed in the region of the ear injected with NanoLuc but not in surrounding noninjected regions of the same ear. No signal was observed in the venous compartment, likely reflecting that blood flow rapidly diluted the bioluminescence signal (fig. S7E). Together, this analysis shows that a small molecule, FFz, can enter the central nervous system (CNS) from the blood and activate an enzyme, NanoLuc, present in CSF, resulting in the generation of photons along the wall of the venous sinus. On the basis of the juxtaposition of SLYM and the venous endothelium in histological examination (Fig. 4A), the selective generation of photons when luciferase was injected into CSF, and the fact that the substrate was present in the vascular compartment (fig. S7, A to C), we propose that the apposition of the venous endothelia and SLYM represents rodent arachnoid villus–like structures, comparable to those in human brain.

The mesothelium surrounding peripheral organs acts as an immune barrier (*26*). Does SLYM also impede the entry of exogenous particles into CSF? In vivo two-photon imaging of Prox1-EGFP$^+$ mice injected intravenously with rhodamine 6G (Rhod6G) to label leukocytes (*35*) showed that a large number of Rhod6G$^+$ myeloid cells are embedded in SLYM (Fig. 5A). The number of Rhod6G$^+$ leukocytes in dura and SLYM was directly comparable, suggesting a prominent role of SLYM in CNS immune responses, which supports the finding that leptomeninges are densely populated with immune cells (*36*) (Fig. 5A). How do systemic inflammation and aging affect the immune cell populations residing in SLYM? Ex vivo analysis of brain sections obtained from Prox1-EGFP$^+$ mice showed that,

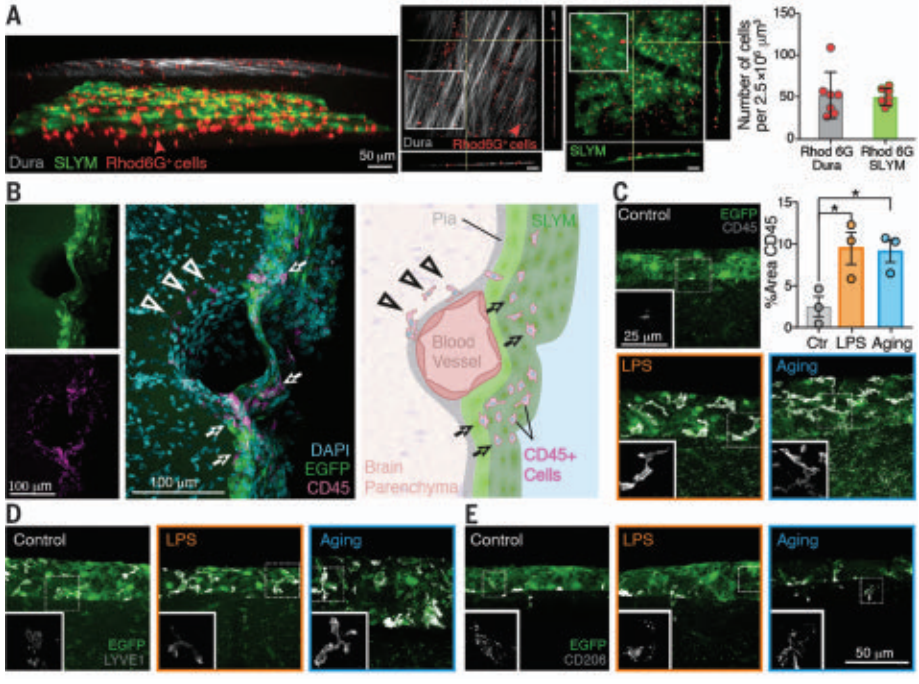


Fig. 5. SLYM hosts a large number of myeloid cells. (**A**) (Left) In vivo two-photon microscopy of Prox1-EGFP$^+$ mice injected with Rhod6G (red) shows that SLYM (EGFP$^+$, green) is permeated by myeloid cells similar to dura (collagen fibers, gray). Middle panels show orthogonal sections depicting Rhod6G $^+$ cells in dura and SLYM, respectively. (Right) In vivo quantification of the number of Rhod6G$^+$ cells present in dura and SLYM. Values are expressed as mean ± SEM, two-tailed unpaired *t* test with Welch's correction, P = 0.5748, n = 7 mice. (**B**) Representative image showing the accumulation of CD45$^+$ cells along the pial vessels. (**C**) The percentage of area covered by CD45$^+$ cells was significantly increased both in aged (12- to 13-month-old) mice and in response to inflammation (LPS 4 mg/kg, ip, 24 hours). Values are expressed as mean ± SEM, two-tailed unpaired *t* test with Welch's correction, $P < 0.005$, n = 3 mice. (**D**) LYVE1 macrophages were also found in the SLYM layer, being more prominent in aged and LPS-treated animals than in healthy young Prox1-EGFP mice. (**E**) The mannose receptor CD206 was detected at similar levels in the young, aged, and LPS-treated groups, suggesting that SLYM may act as a niche for border-associated mouse macrophages. Significance shown as *$P < 0.05$. Ctr, control.

in the control group, $CD45^+$ cells were abundant, located mostly along pial vessels in the surface of the brain (Fig. 5B). This observation, together with the significant increase in $CD45^+$ in inflammation-prone conditions [aging and lipopolysaccharide (LPS)–treated mice, 4 mg/kg of body weight, intraperitoneally (ip), 24 hours] (Fig. 5C), suggests that SLYM can act as a $CD45^+$ recruiting and/or proliferating site in pathological conditions. Of note, the dose of LPS used (4 mg/kg) did not affect the BBB (fig. S8). Additional immune markers showed that $LYVE1^+$ (Fig. 5D), $CD206^+$ (Fig. 5E), and $CD68^+$ (fig. S9) macrophages can be found in SLYM, together with dendritic cells ($CD11c^+$) (fig. S9). Despite the absence of $CD3^+$ and $CD19^+$ lymphocytes (fig. S9), our results indicate that SLYM functions as a niche for immunological surveillance. Thus, in young, healthy mice, SLYM hosts $CD45^+$ cells, but the number and diversity of innate immune cells rapidly expands in LPS-induced inflammation and was also significantly altered in aged mice. We conclude that SLYM fulfills the characteristics of a mesothelium by acting as an immune barrier that prevents exchange of small solutes between the outer and inner subarachnoid space compartments and by covering blood vessels in the subarachnoid space.

Discussion

The critical roles of the meningeal membranes lining the brain have only recently been acknowledged (*5*, *37*). It is now known that CSF is drained by a network of lymphatic vessels in the meninges and that suppression of this drainage accelerates protein aggregation and cognitive decline in animal models of neurodegeneration (*38*–*40*). SLYM is $Prox1^+$/$PDPN^+$/$LYVE1^-$/$CRABP2^+$/$VEGFR3^-$/$CLDN\text{-}11^-$/$E\text{-}Cad^-$ and thereby distinct from the traditional meningeal membranes, including dura, arachnoid, and pia, as well as the meningeal lymphatic vessels and the arachnoid trabecula. SLYM subdivides the subarachnoid space into two compartments, suggesting that CSF transport is more organized than currently acknowledged. For example, SLYM covering the vasculature in the inner subarachnoid space will guide CSF influx along the penetrating arterioles into the brain parenchyma without circulating solutes present in the outer subarachnoid space compartment. Yet the discovery of a fourth meningeal layer, SLYM, has several implications beyond fluid transport. The observation that SLYM is a barrier for CSF solutes that have a molecular weight larger than 3 kDa will require more detailed studies but indicates a need to redefine the concept of CNS barriers to include SLYM. The meningeal membranes are hosts to myeloid cells responsible for immune surveillance of the CNS (*5*, *37*), and SLYM, owing to its close association with the brain surfaces, is likely to play a prominent role in this surveillance. Herein, we showed a large increase in the number and diversity of immune cells residing in SLYM in response to acute inflammation and natural aging. Physical rupture of SLYM could, by altering CSF flow patterns, explain the prolonged suppression of glymphatic flow after traumatic brain injury as well as the heightened posttraumatic risk of developing Alzheimer's disease (*41*, *42*). Rupture of SLYM will also permit the direct passage of immune cells from the skull bone marrow (*33*, *43*) into the inner subarachnoid space, with direct access to the brain surfaces, possibly explaining the prolonged neuroinflammation after traumatic brain injury (*44*). SLYM may also be directly involved in CNS immunity, in addition to being host to many immune cells. Lymphatic-like tissues can transform quickly in the setting of inflammation, which in the brain may be of notable relevance for diseases such as multiple sclerosis (*45*).

REFERENCES AND NOTES

1. J. J. Iliff *et al*., *Sci. Transl. Med.* **4**, 147ra111 (2012).
2. H. Mestre *et al*., *Nat. Commun.* **9**, 4878 (2018).
3. N. E. Fultz *et al*., *Science* **366**, 628–631 (2019).
4. H. Mestre *et al*., *eLife* **7**, e40070 (2018).
5. A. Louveau *et al*., *Nature* **523**, 337–341 (2015).
6. A. Aspelund *et al*., *J. Exp. Med.* **212**, 991–999 (2015).
7. L. H. Weed, *J. Anat.* **72**, 181–215 (1938).
8. A. Key, G. Retzius, *Studien in der Anatomie des Nervensystems und des Bindegewebes* (Samson & Wallin, 1876).
9. W. E. le Gros Clark, *J. Anat.* **55**, 40–48 (1920).
10. K. Ohta *et al*., *Kurume Med. J.* **49**, 177–183 (2002).
11. M. K. Rasmussen, H. Mestre, M. Nedergaard, *Physiol. Rev.* **102**, 1025–1151 (2022).
12. A. Drieu *et al*., *Nature* **611**, 585–593 (2022).
13. I. Choi *et al*., *Blood* **117**, 362–365 (2011).
14. J. T. Wigle *et al*., *EMBO J.* **21**, 1505–1513 (2002).
15. K. Masamoto *et al*., *Neuroscience* **212**, 190–200 (2012).
16. A. Nimmerjahn, F. Kirchhoff, J. N. Kerr, F. Helmchen, *Nat. Methods* **1**, 31–37 (2004).
17. K. Kothur, L. Wienholt, F. Brilot, R. C. Dale, *Cytokine* **77**, 227–237 (2016).
18. H. Mestre, Y. Mori, M. Nedergaard, *Trends Neurosci.* **43**, 458–466 (2020).
19. M. Tomooka, C. Kaji, H. Kojima, Y. Sawa, *Acta Histochem. Cytochem.* **46**, 171–177 (2013).
20. S. Banerji *et al*., *J. Cell Biol.* **144**, 789–801 (1999).
21. M. A. Asson-Batres, O. Ahmad, W. B. Smith, *Cell Tissue Res.* **312**, 9–19 (2003).
22. C. B. Brøchner, C. B. Holst, K. Møllgård, *Front. Neurosci.* **9**, 75 (2015).
23. J. DeSisto *et al*., *Dev. Cell* **54**, 43–59.e4 (2020).
24. J. Derk, H. E. Jones, C. Como, B. Pawlikowski, J. A. Siegenthaler, *Front. Cell. Neurosci.* **15**, 703944 (2021).
25. M. M. Mortazavi *et al*., *World Neurosurg.* **111**, 279–290 (2018).
26. S. E. Mutsaers, F. J. Pixley, C. M. Prêle, G. F. Hoyne, *Curr. Opin. Immunol.* **64**, 88–109 (2020).
27. B. A. Hills, J. R. Burke, K. Thomas, *Perit. Dial. Int.* **18**, 157–165 (1998).
28. A. Jayatilaka, *Ceylon J. Med. Sci.* **18**, 25–30 (1969).
29. D. G. Potts, V. Deonarine, *J. Neurosurg.* **38**, 722–728 (1973).
30. K. Suzuki *et al*., *Nat. Commun.* **7**, 13718 (2016).
31. Y. Su *et al*., *Nat. Methods* **17**, 852–860 (2020).
32. M. P. Hall *et al*., *ACS Chem. Biol.* **7**, 1848–1857 (2012).
33. J. Rustenhoven *et al*., *Cell* **184**, 1000–1016.e27 (2021).
34. P. Mastorakos, D. McGavern, *Sci. Immunol.* **4**, eaav0492 (2019).
35. H. Baatz, M. Steinbauer, A. G. Harris, F. Krombach, *Int. J. Microcirc. Clin. Exp.* **15**, 85–91 (1995).
36. A. Merlini *et al*., *Nat. Neurosci.* **25**, 887–899 (2022).
37. A. Louveau *et al*., *J. Clin. Invest.* **127**, 3210–3219 (2017).
38. S. Da Mesquita, Z. Fu, J. Kipnis, *Neuron* **100**, 375–388 (2018).
39. Z. Xu *et al*., *Mol. Neurodegener.* **10**, 58 (2015).
40. L. Wang *et al*., *Brain Pathol.* **29**, 176–192 (2019).
41. A. Z. Mohamed, P. Cumming, J. Götz, F. Nasrallah; Department of Defense Alzheimer's Disease Neuroimaging Initiative, *Eur. J. Nucl. Med. Mol. Imaging* **46**, 1139–1151 (2019).
42. J. D. Flatt, P. Gilsanz, C. P. Quesenberry Jr., K. B. Albers, R. A. Whitmer, *Alzheimers Dement.* **14**, 28–34 (2018).
43. S. Brioschi *et al*., *Science* **373**, eabf9277 (2021).
44. J. J. Iliff *et al*., *J. Neurosci.* **34**, 16180–16193 (2014).
45. N. B. Pikor, A. Prat, A. Bar-Or, J. L. Gommerman, *Front. Immunol.* **6**, 657 (2016).

ACKNOWLEDGMENTS

We acknowledge P. S. Froh and H. Nguyen (Department of Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark) for their excellent technical assistance for the histology and immunohistochemistry of the decalcified samples. We also thank D. Xue for expert graphical support, B. Sigurdsson for analysis, and H. Hirase, N. Cankar, and N. C. Petersen for critical reading of the manuscript. **Funding:** Funding was provided by Lundbeck Foundation grant R386-2021-165 (M.N.), Novo Nordisk Foundation grant NNF20OC0066419 (M.N.), the Vera & Carl Johan Michaelsen's Legat Foundation (K.M.), National Institutes of Health grant R01AT011439 (M.N.), National Institutes of Health grant U19NS128613 (M.N.), US Army Research Office grant MURI W911NF1910280 (M.N.), Human Frontier Science Program grant RGP0036 (M.N.), the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (M.N.), and Simons Foundation grant 811237 (M.N.). The views and conclusions contained in this article are solely those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the National Institutes of Health, the Army Research Office, or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. The funding agencies have taken no part on the design of the study, data collection, analysis, interpretation, or in writing of the manuscript. **Author contributions:** K.M. and M.N. designed the study. F.R.M.B., P.K., L.M.M., C.D., V.P., M.K.R., R.S.G., N.L.H., T.E., and Y.M. performed the experiments, collected the data, and performed the analysis. K.M. and M.N. wrote the manuscript. All authors read and approved the final version of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adc8810
Materials and Methods
Figs. S1 to S9
Table S1
References (*46*–*54*)
MDAR Reproducibility Checklist
Movie S1

Submitted 6 May 2022; resubmitted 13 September 2022
Accepted 7 December 2022
10.1126/science.adc8810

METALLURGY

Machine learning–aided real-time detection of keyhole pore generation in laser powder bed fusion

Zhongshu Ren[1], Lin Gao[1], Samuel J. Clark[2], Kamel Fezzaa[2], Pavel Shevchenko[2], Ann Choi[3,4], Wes Everhart[3], Anthony D. Rollett[4], Lianyi Chen[5], Tao Sun[1]*

Porosity defects are currently a major factor that hinders the widespread adoption of laser-based metal additive manufacturing technologies. One common porosity occurs when an unstable vapor depression zone (keyhole) forms because of excess laser energy input. With simultaneous high-speed synchrotron x-ray imaging and thermal imaging, coupled with multiphysics simulations, we discovered two types of keyhole oscillation in laser powder bed fusion of Ti-6Al-4V. Amplifying this understanding with machine learning, we developed an approach for detecting the stochastic keyhole porosity generation events with submillisecond temporal resolution and near-perfect prediction rate. The highly accurate data labeling enabled by operando x-ray imaging allowed us to demonstrate a facile and practical way to adopt our approach in commercial systems.

After more than three decades of intense research and development, laser powder bed fusion (LPBF) has advanced from a convenient rapid prototyping tool for shortening the design cycle toward a manufacturing technology for producing end-use metallic components (*1*–*3*). Although some industries fully embrace LPBF now, others are more cautious about quality control as they integrate it into their product lines. As a primary metal additive manufacturing (AM) technique, LPBF is capable of fabricating parts with complex geometries and fine features. However, some technical barriers still need to be overcome before LPBF can reach its full potential as a disruptive manufacturing technique. In a typical LPBF process, a high-power laser beam is used to locally melt and consolidate metal powder to form three-dimensional (3D) objects layer by layer. The extreme thermal conditions involved in the printing process trigger transient phenomena and complex structural dynamics. Their interplay often leads to structural defects, such as porosity. One common porosity is caused by the momentary collapse of the vapor depression zone, known as keyhole porosity (*4*–*8*).

Under the condition of excess laser energy input (high power and slow scan velocity), metal vaporization exerts a recoil pressure that pushes down the melt pool surface, forming a narrow and deep keyhole in which multiple laser reflection and absorption events occur (*9*–*11*). Although this increases the overall laser absorption by the metal and that benefits the manufacturing process by improving the energy efficiency and increasing the build rate, the nonuniform laser absorption on the keyhole wall generates local hotspots and causes imbalance between the recoil pressure, vapor dynamic pressure, capillary force, and Marangoni force. Under unstable keyhole conditions, gas bubbles pinch off the keyhole tip, and some eventually become pore defects when they are trapped by the advancing solidification front (*4*, *5*, *12*). For LPBF of a given material, the unstable keyhole zone in the power-velocity (P-V) process map can be defined with small uncertainty (*4*, *5*, *12*). Setting the initial laser parameters outside the unstable keyhole zone helps mitigate the generation of keyhole porosity. However, multiple factors involved in LPBF can still offset the laser melting mode and create keyhole-porosity–prone conditions, such as the drift of laser spot size, power, and scan speed, as well as the scan strategies that result in local overheating. Therefore, keyhole porosity may still exist in a part even if the initial machine setting is optimized for printing certain material.

Real-time detection of keyhole pore generation in LPBF is critical not only for facilitating post-build part interrogation and qualification but also for developing closed-loop control systems that can anticipate the need for local variations during the build process. Optical and acoustic sensors are commonly used as process monitors, and data analysis approaches have been developed to correlate the process signatures with porosity (*13*–*15*). Although some successes were achieved by the community, particularly with application of machine learning (*16*–*18*), for differentiating pore-prone conditions from normal conditions, detection of the generation of a keyhole pore locally and instantaneously remains challenging. This problem is because the stochastic nature of keyhole pore generation has other physics roots, beyond the apparent cause of processing parameter drifting. During a laser scan with an unstable keyhole condition, the exact locations where gas bubbles form are random, and which bubbles will become pore defects eventually and which will be recaptured by the keyhole and vanish is not deterministic.

We discovered two modes of keyhole oscillation in Ti-6Al-4V under unstable keyhole conditions using simultaneous high-speed synchrotron x-ray and thermal imaging. We then developed a methodology for detecting keyhole pore generation by integrating experimental data, multiphysics simulation, and machine learning. We used the thermal signals emitted from the keyhole region for predicting the pore-generation events. X-ray images of the LPBF process provide a data-rich ground truth for calibrating and validating the theoretical model and for training the machine-learning algorithm. We achieved near-perfect detection accuracy with submillisecond temporal resolution for both powder bed and bare substrate samples.

Operando experiment and data analytics

We conducted our operando experiment at the 32-ID-B beamline of the Advanced Photon Source of Argonne National Laboratory (Fig. 1A). High-energy x-rays passed through the single-layer powder bed or bare substrate samples to reveal the subsurface structure dynamics, and we set up a thermal camera to collect an angled top view of the melt pool. A continuous-wave–mode fiber laser with a Gaussian profile and wavelength of 1070 nm scanned the sample along a single straight line at various powers and speeds. We collected full-field x-ray images (Fig. 1C) from single-track melting events with a spatial resolution of 2 to 3 μm/pixel, temporal resolution of 0.1 ns to 7.5 μs, and frame rates of 50 kHz to 1.08 MHz. Simultaneously, we collected thermal images (Fig. 1B) of the melt track in the spectra range of visible light to infrared, view angle of 38° to 58°, spatial resolution of 5 to 30 μm/pixel, temporal resolution of 0.3 to 5 μs, and frame rates of 50 to 200 kHz [(*19*), sections 1.2 to 1.4].

By setting a threshold, we extracted the average light emission intensity from the keyhole region, which condensed the 2D optical images into 1D plots (Fig. 1D). Such data compression can alleviate the burden of computer storage substantially in real-time monitoring because the high-resolution images can be immediately processed with on-camera-chip. Meanwhile, 1D time-series signals facilitate identification of frequency-related features associated with keyhole dynamics. We performed wavelet analyses over the segmented 1D datasets to create scalograms (Fig. 1E), which reveal the characteristic oscillations localized in the time domain. We fed short-window scalograms, labeled by operando

[1]Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA 22904, USA. [2]X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Lemont, IL 60439, USA. [3]Kansas City National Security Campus Managed by Honeywell Federal Manufacturing and Technologies, US Department of Energy, Kansas City, MO 64147, USA. [4]Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [5]Department of Mechanical Engineering, University of Wisconsin–Madison, Madison, WI 53706, USA.
*Corresponding author. Email: ts7qw@virginia.edu

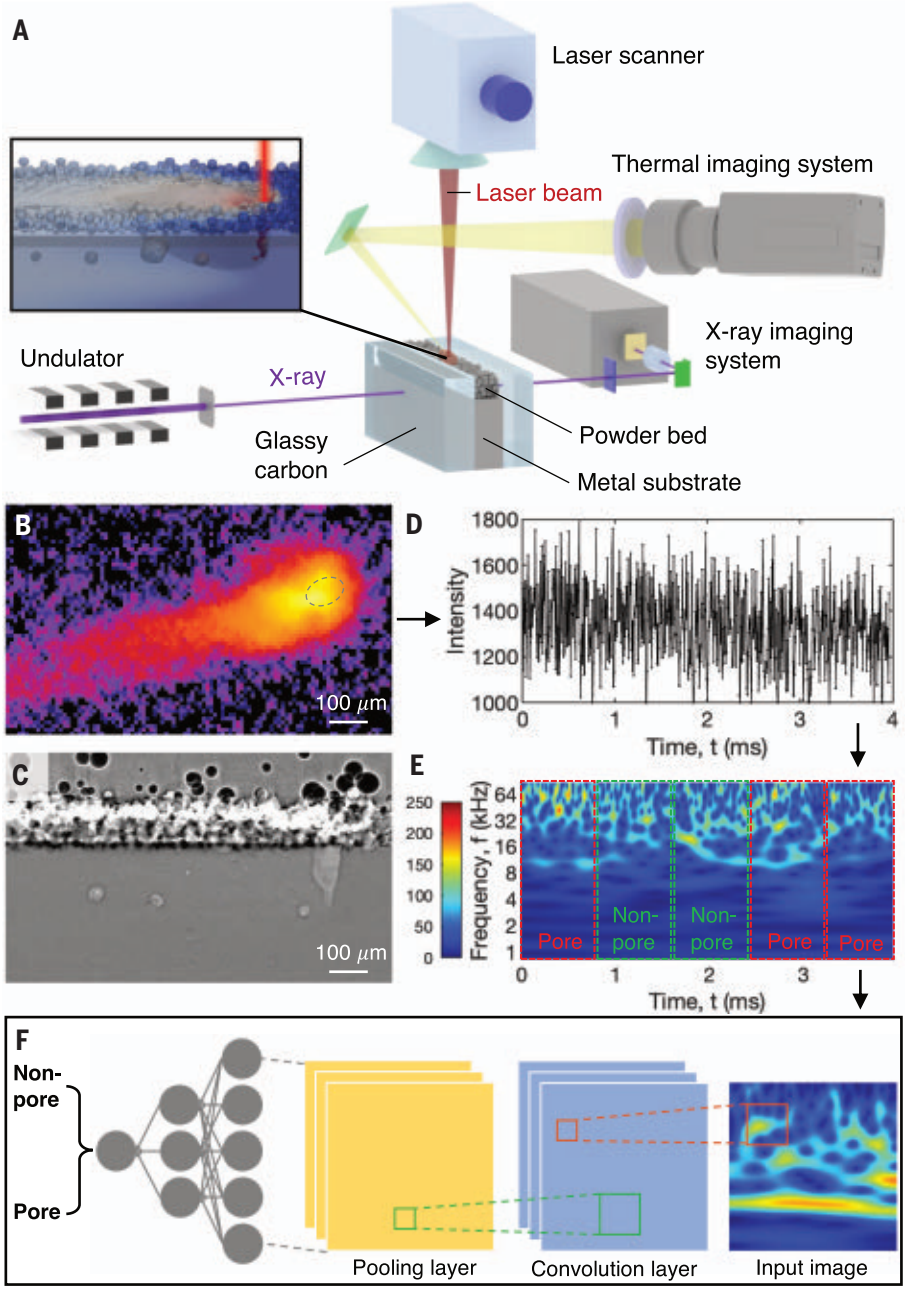


Fig. 1. Real-time keyhole porosity detection in LPBF. (**A**) Schematic of the simultaneous synchrotron x-ray and thermal imaging experiment on scanning laser melting of Ti-6Al-4V. (**B**) A representative angle top-view thermal image. (**C**) A representative side-view x-ray image. (**D**) Typical time-series signal of the average emission intensity from the keyhole region [(B), dashed oval] extracted from the thermal image sequence. (**E**) Wavelet analysis performed over the time-series signals in (D). The scalogram is sectioned into a few windows, which are then labeled as either "Non-pore" or "Pore" on the basis of the operando x-ray imaging result. (**F**) Machine-learning approach with sectioned scalograms as input data. A CNN was used, which is composed of a series of alternating convolution and pooling layers and a final layer. Each convolution layer extracts features from its previous layer, using filters learned from the trained model, to form a feature map. The feature map is then down-sampled by a pooling layer to reduce the number of parameters to learn. The final layer of the CNN classifies the input scalogram as either "Non-pore" or "Pore."

x-ray imaging results as cases of "Pore" and "Non-pore," into a convolutional neural network (CNN) for predicting the keyhole pore-generation events (Fig. 1F and fig. S3) [(*19*), section 1.6].

Intrinsic keyhole oscillation

Our approach for detecting keyhole pore generation by using thermal imaging was developed on the basis of our observation and understanding of keyhole oscillations. In this work, we define the two modes as intrinsic and perturbative oscillations. The intrinsic oscillation occurs for both stable and unstable keyholes (Fig. 2, A and C to F; figs. S6 to S12; and movies S1 to S9) and is a consequence of varying balance between Marangoni force, surface tension, and recoil pressure. In one oscillation cycle (Fig. 2, C to F; figs. S6 to S12; and movies S1 to S9), when the keyhole has the smallest opening at the rim, the rear keyhole wall is directly exposed to the incident laser beam and heated to a high temperature (Fig. 2C, 15 μs, and figs. S8 to S12, 0 μs). The Marangoni force then convects the hot liquid (the hump structure near the rear keyhole rim shown in Fig. 2D, 40 μs, and figs. S9, 5 μs; S10, 20 and 60 μs; and S12, 22 μs) back toward the rear of melt pool, where the temperature is lower. The liquid flow at the melt pool surface expands the opening of the keyhole and cools it. When most of the hot liquid flows backward [the hump structure is now at the rear location (Fig. 2D, 60 μs, and fig. S12, 54 μs)], the keyhole temperature drops to its minimum (Fig. 2C, 90 μs, and fig. S8, 14 μs), and the opening reaches its maximum. After this point, the Marangoni force becomes negligible owing to the much reduced temperature gradient around the keyhole. Instead, the surface tension starts to dominate the keyhole dynamics and drives the liquid in the rear keyhole wall to flow forward and close up the keyhole. As the keyhole opening shrinks, more liquid is exposed to the incident laser beam, and the thermal gradient in the keyhole region increases. The Marangoni force starts to drive the liquid flow again, initiating a new cycle of intrinsic oscillation [(*19*), section 2.4]. The type of keyhole oscillation is consistent with those observed in the previous studies (*13*, *20*–*23*).

We used the thermal information of the keyhole collected by our near-infrared (NIR) camera as an indicator of the intrinsic oscillation (representative intrinsic oscillation is shown in in Fig. 2, E and F, 16 kHz). The mechanism for the intrinsic oscillation of an unstable keyhole conditions is the same as that of a stable keyhole (figs. S9 to S12 and movies S5 to S9). The intrinsic oscillation is not responsible for the keyhole pore generation under unstable keyhole conditions. This is because even though there may be temporary keyhole collapse and bubble generation, the potent forward liquid flow involved in the intrinsic oscillation can push the bubble toward the advancing laser and allow it to be captured by the rapidly drilling keyhole (Fig. 2A, figs. S11 and S12, and movies S8 and S9).

Perturbative keyhole oscillation

The perturbative oscillation only occurs under unstable keyhole conditions (Fig. 2, B and G to J; figs. S13 to S16; and movies S10 to S14). The "perturbative" term is because the associated

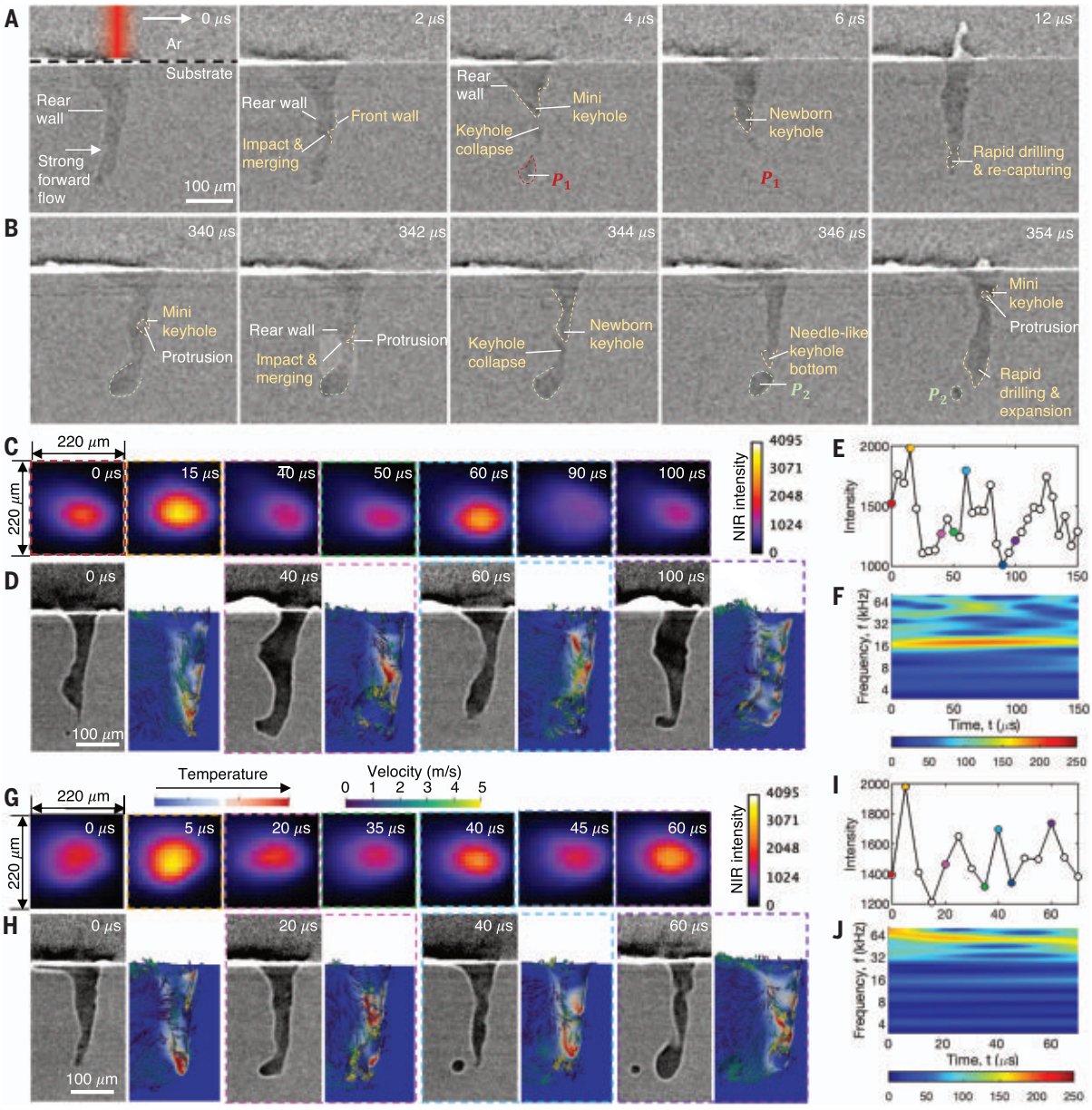


Fig. 2. Intrinsic and perturbative keyhole oscillations in Ti-6Al-4V. (**A**) Megahertz x-ray images of intrinsic keyhole oscillation with no keyhole pore generated. (**B**) Megahertz x-ray images of perturbative keyhole oscillation with a keyhole pore generated. (**C**) Thermal images of intrinsic keyhole oscillation. (**D**) High-speed x-ray images and corresponding multiphysics simulation of intrinsic keyhole oscillation. (**E**) Time-series signal of the average light emission intensity around the keyhole, extracted from the thermal image sequence in (C). (**F**) Scalogram corresponding to the time-series signals in (E). (**G**) Thermal images of perturbative keyhole oscillation. (**H**) High-speed x-ray images and corresponding multiphysics simulation of perturbative keyhole oscillation. (**I**) Time-series signal of the average light emission intensity around the keyhole, extracted from the thermal image sequence in (G). (**J**) Scalogram corresponding to the time-series signals in (I). The laser power and speed are 200 W and 400 mm/s, respectively. The color markers in (E) and (I) indicate data points with corresponding frames shown in (C), (D), (G), and (H), which have borders with the same colors.

keyhole behavior disrupts the intrinsic oscillation. An unstable keyhole features highly dynamic protrusion structures on the keyhole wall (*4*, *22*, *24*). A protrusion on the front keyhole wall has two effects. First, the protrusion shadows the liquid right underneath it from being heated by the incident laser directly. Thus, a large temperature difference between the liquid above and below the protrusion develops, creating complex flow patterns. Toward the keyhole opening, the Marangoni force drives the hot liquid moving upward and backward, which opens up the keyhole rim (Fig. 2H, simulation at 20, 40, and 60 μs; and fig. S15, 0 μs) (*5*). Second, the protrusion reflects the incident laser ray toward the middle and upper regions of the rear keyhole wall, causing local hotspots and consequently a decrease of recoil pressure at the keyhole bottom. Both of these effects increase the chance of keyhole wall collapsing (Fig. 2B, 342 and 344 μs, and figs. S13, 0 and 2 μs, and S16, 2 and 4 μs). Once the collapsing occurs, the new keyhole (upper part of the previous keyhole) is directly exposed to the laser, and this keyhole's temperature increases (Fig. 2G, 40 μs, and fig. S15, 22.5 μs). The high recoil pressure

then pushes down the keyhole rapidly (*22*, *24*). Meanwhile, the Marangoni force transports the hot liquid from the bottom of the newborn keyhole toward the rear rim, facilitating the drilling process. During this process, a new protrusion appears on the front keyhole wall, and a fresh cycle of perturbative oscillation starts (Fig. 2, B, 354 μs, and H, 60 μs; and figs. S13, 6 μs; S15, 5 and 17.5 μs; and S16, 8 μs). The frequency of perturbative oscillation is generally above 40 kHz (figs. S14 and S15), which is higher than the intrinsic oscillation frequency. Owing to the chaotic nature of an unstable keyhole and the random occurrence of protrusions on front and rear keyhole walls, the perturbative oscillation frequency varies dramatically even during a single laser scan.

We discovered that the high-frequency perturbative oscillation is responsible for the keyhole porosity generation in Ti-6Al-4V. Once the keyhole collapses, the lower part of the keyhole is separated from the upper part, appearing as a gas bubble pinching off the keyhole tip. The bubble is then pushed downward and backward by the new keyhole (Fig. 2, B, 346 to 354 μs, and H, 40 to 60 μs; and figs. S13, 12 μs; S14, 20 μs; S15, 10 μs; and S16, 14 μs) (*4*). As the bubble moves away from the keyhole, its motion is primarily controlled by the fluid flow in the melt pool. Very few pores can escape by flowing to the melt pool surface or circulating back to the keyhole, whereas the majority are trapped by the advancing solidification front and become pore defects (*4*). The perturbative keyhole oscillation is the key to answering a question that has not been asked previously: when printing by using a single P-V condition in unstable keyhole mode, why some bubbles can escape the keyhole and become pores, whereas others are recaptured by the keyhole and disappear [(*19*), sections 2.4 to 2.6].

Machine learning aided keyhole pore detection

The distinct frequencies of intrinsic and perturbative oscillations allowed us to detect the keyhole pore–generation events with high fidelity and high resolution. We divided each scalogram of thermal data we obtained from a single line scan into several segments, which correspond to frequency patterns within a smaller time window. We used the labeled scalogram segments to train a CNN algorithm for classifying the test data into "Pore" and "Non-pore" cases. We used a feature visualization technique known as gradient-weighted class activation mapping (Grad CAM) (*25*) to highlight the important features corresponding to these two categories. The machine-learning algorithm used the low-frequency feature (intrinsic oscillation) to classify the "Non-pore" cases and the high-frequency feature (perturbative oscillation) to classify the "Pore" cases (fig. S21).

After confirming that our machine-learning model is capable of identifying features with predictive power from the input scalograms, we investigated other factors that may affect the keyhole porosity prediction rate. These include the configuration of thermal imaging system (table S2), spatial resolution (table S3), frame rate (table S3 and fig. S24), view angle (the angle between optical axis and sample surface) (fig. S25 and table S4), and the window length of sectioned scalograms (figs. S24 and S25). Because wavelet analysis is essentially a form of short-window Fourier transform, increasing the frame rate of raw data and the window length of each scalogram can both increase the information a scalogram reveals (fig. S24) and thereby improve the keyhole pore prediction rate. Specifically, because the frequency of perturbative oscillation is higher than 40 kHz, the thermal imaging rate needs to be above 80 kHz according to the Nyquist-Shannon sampling theorem. Therefore, raw data with 50 kHz rate, regardless of the optical configuration, yielded much reduced prediction rate (tables S2 to S4).

In our approach, the off-axial thermal imaging is preferred than the commonly used coaxial setup, because it (i) offers flexibility and ease of installation to existing 3D printers, whereas the coaxial setup involves complicated optical systems (*26*–*28*); (ii) provides more reliable process signals than those of the coaxial setup, in which the thermal emission travels back through optics such as the F-Theta lens and dielectric mirror to the imager, influencing the signals (*29*, *30*); and (iii) an optimum angle of ~50° was found in our simulation (fig. S25) that yielded the highest keyhole pore prediction rate. We speculate that the reason for the lower prediction rate from coaxial detection is twofold. First, when looking straight down, the camera collects the thermal information of the keyhole bottom, which remains hot most of the time, so it may be less sensitive to the keyhole oscillation [more discussion is available in (*19*), section 2.11]. Second, an unstable keyhole appears at relatively slow laser scanning speeds. Under these conditions, the vapor plume ejects upward with only small inclination angles. The fluctuation of the vapor plume reflects the keyhole oscillation to a great extent, but at some distance above the sample when the concentration and speed of the vapor plume diminish, its dynamic motion is also driven by the ambient gas flow. Such vapor plume effect will likely weaken when introducing shielding gas flow across the powder bed, and many commercial LPBF machines offer this function now. By contrast, at ~50° view angle, although the effective spatial resolution is slightly decreased, the camera sees only the upper part of the keyhole wall, with limited influence of the secondary vapor plume fluctuation. The effect of view angle was partially proved by our experimental result (table S4), yet the exact vapor plume effect on thermal imaging in different geometries demands more sophisticated experimental and modeling efforts [more discussion is available in (*19*), section 2.7].

For powder bed samples, 100% prediction rate (including "accuracy," "precision," and "recall" results) of keyhole pore generation in Ti-6Al-4V was achieved by performing NIR imaging at 48° view angle, 200 kHz sampling rate, 10 μm spatial resolution, and 0.75 ms time window (table S5). We demonstrated that the stochastic keyhole pore occurrence can be truly detected locally, among laser scans under various P-V conditions covering all melting modes, or even the same P-V condition in unstable keyhole mode, in which the number of pores are smaller than the keyhole-collapsing and bubble-forming events.

Implementation without additional synchrotron experiments

A key factor for achieving high prediction rate by using the machine-learning approach is the accurate data labeling. The larger the mislabeling rate is, the worse the performance of the machine-learning algorithm (Fig. 3). In previous efforts (*17*, *18*, *31*), porosity in the printed samples was typically characterized by using x-ray computed tomography. Such an ex situ (post mortem) data-labeling approach identifies the final pore locations rather than the moment the bubbles were initially generated. Because a bubble can move within the melt pool after it pinches off the keyhole, a substantial mislabeling rate results from ex situ labeling when the window length is small (Fig. 3E). In our operando synchrotron x-ray imaging experiments, we could measure not only the final location of a keyhole pore but also the exact moment at which the bubble forms (Fig. 3, A and B). We performed statistical analysis over the experiment and simulation data on powder bed and plate samples (Fig. 3C). The line regression shows that ex situ labeling could induce an uncertainty of the pore generation time by ~0.4 ms, which caused considerable reduction in prediction rate (Fig. 3F).

With the calibrated simulation, optimized thermal imaging scheme and machine-learning algorithm, we demonstrate two practical means to effectively apply our approach without the need of additional synchrotron experiments. The first approach is to calibrate the ex situ–labeled time by adding 0.4 ms to better represent the moments when keyhole pores form (bubble separating from the keyhole tip). This time offset appears to be independent of laser power and scan speed but may vary with the laser spot size and material. The calibration of time stamps served to reduce the mislabeling rate (Fig. 3E) and improve the prediction accuracy substantially (Fig. 3F). Encouraged by

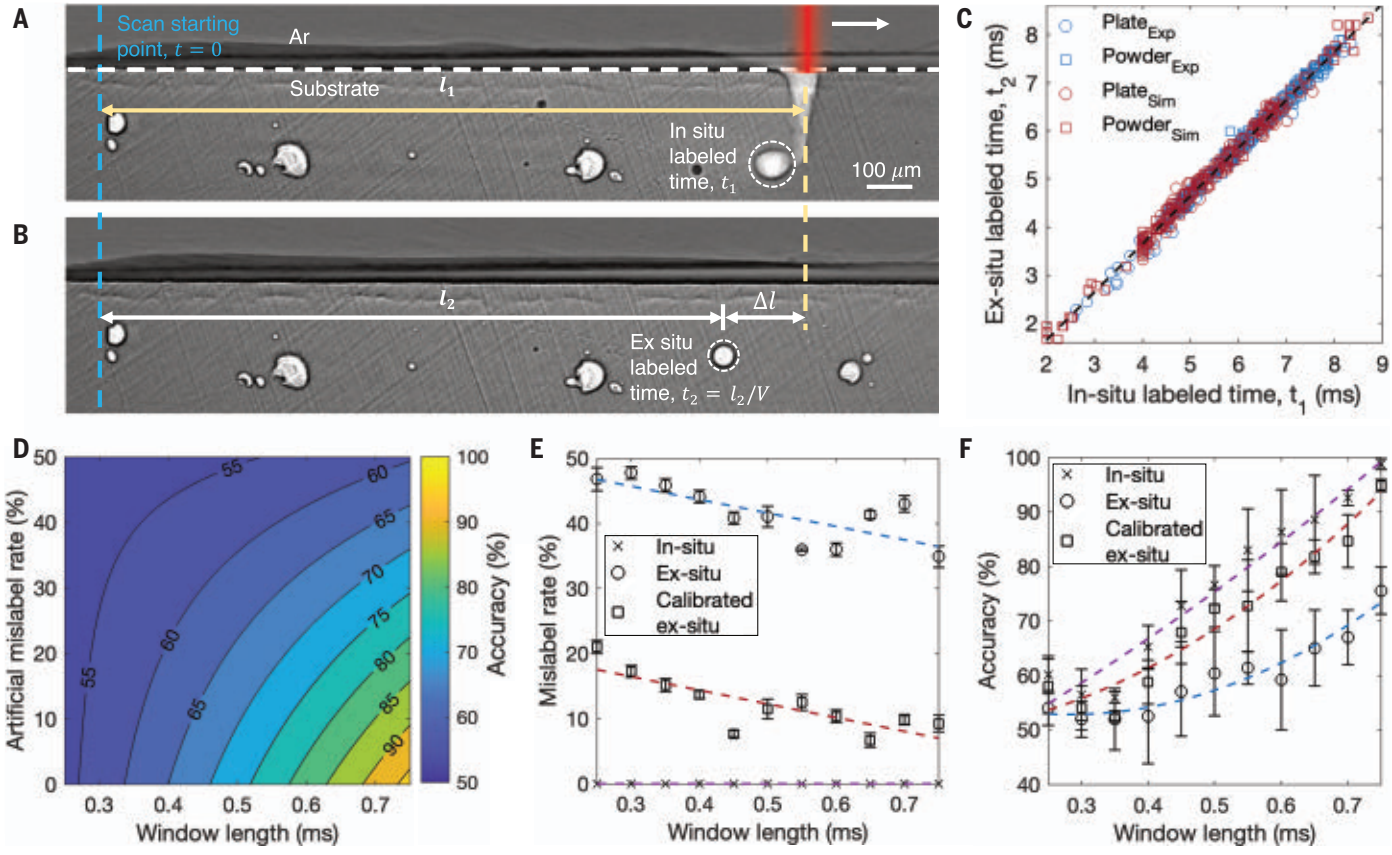


Fig. 3. Effect of data mislabeling on prediction rate. (**A**) Illustration of in situ–labeled time t_1 of a keyhole pore, the moment when a bubble pinches off the keyhole. (**B**) Illustration of ex situ–labeled time t_2 of the same pore as marked in (A), an artificial moment corresponding to the final pore location. (**C**) Ex situ–labeled time t_2 as a function of in situ–labeled time t_1 for experimental and simulation data on powder bed and plate samples. A linear regression is indicated with the black dashed line. The error bars are not displayed in this plot because the uncertainties in both t_1 and t_2 (~0.02 ms) are much smaller than the time scale under consideration. (**D**) Contour map of machine-learning prediction accuracy as functions of mislabeled rate and scalogram window length. The variable of mislabeled rate is created by randomly assigning a certain percentage of the scalogram the inverse label of the in situ labels in the training data. (**E**) Mislabeled rate as a function of scalogram window length for in situ, ex situ, and calibrated ex situ labeling. (**F**) The corresponding machine-learning prediction accuracy as a function of scalogram window length. The error bars in (E) and (F) are the standard deviation of 10 repeated training, representing the randomness in partitioning the training data into training and validation sets.

this success, we performed a proof-of-concept experiment using a commercial LPBF system (SLM Solutions 125) (fig. S26). The size of the chamber and the view window location limited the setting of thermal imaging. Even though we could not apply the optimum imaging condition, the prediction accuracy, precision, and recall all improved after adjusting the ex situ labels (table S6). The second approach was to train the machine-learning model by use of well-calibrated simulations. When tested on the experiment data, prediction accuracies of 85% (0.75 ms window) and 87% (1.5 ms window) were obtained for the plate and powder bed samples, respectively (table S7). It was notably more difficult to predict the pore generation in powder bed samples by using such an approach. This is different than the case of using experimental data for training, in which the prediction rate is generally higher for powder bed samples. The discrepancy highlights the need for developing more sophisticated LPBF models, in which the physics governing the powder motion and vapor plume dynamics should be considered.

Concluding remarks

The research outcome we describe highlights the enabling character of operando synchrotron x-ray imaging experiments. It not only provided key information to calibrate the multiphysics model but also allowed us to discover the distinctive keyhole oscillation behavior associated with keyhole porosity. The quantitative understanding built the foundation for our machine-learning model that detects keyhole pore formation from thermal imaging. With the ground truth acquired with operando x-ray imaging, our approach is capable of detecting the generation of keyhole pores in powder bed samples with 100% accuracy and submillisecond temporal resolution. For those who have limited or no access to synchrotron, this approach can be readily implemented to a LPBF system in two ways as just described. The off-axis detection scheme and the need for only 1D data series could be particularly convenient for existing LPBF systems without preinstalled co-axial imaging optics or high-performance computing hardware. The strategy of detecting build anomalies by examining the oscillation behaviors of the keyhole and melt pool is generic and practical. We believe that the process monitoring systems built on this core concept will promote the qualification and certification of metal AM parts.

REFERENCES AND NOTES

1. T. DebRoy, T. Mukherjee, H. L. Wei, J. W. Elmer, J. O. Milewski, *Nat. Rev. Mater.* **6**, 48–68 (2021).
2. T. Sun, W. Tan, L. Chen, A. Rollett, *MRS Bull.* **45**, 927–933 (2020).
3. T. DebRoy *et al.*, *Nat. Mater.* **18**, 1026–1032 (2019).
4. C. Zhao *et al.*, *Science* **370**, 1080–1086 (2020).
5. M. Bayat *et al.*, *Addit. Manuf.* **30**, 100835 (2019).
6. A. A. Martin *et al.*, *Nat. Commun.* **10**, 1987 (2019).
7. Y. Huang *et al.*, *Nat. Commun.* **13**, 1170 (2022).
8. Z. Gan *et al.*, *Nat. Commun.* **12**, 2379 (2021).
9. S. A. Khairallah *et al.*, *Science* **368**, 660–665 (2020).

10. R. Cunningham *et al.*, *Science* **363**, 849–852 (2019).
11. W. E. King *et al.*, *J. Mater. Process. Technol.* **214**, 2915–2925 (2014).
12. S. A. Khairallah, A. T. Anderson, A. Rubenchik, W. E. King, *Acta Mater.* **108**, 36–45 (2016).
13. S. A. Khairallah, T. Sun, B. J. Simonds, *Additive Manufacturing Letters* **1**, 100002 (2021).
14. J. R. Tempelman *et al.*, *Addit. Manuf.* **55**, 102735 (2022).
15. R. Drissi-Daoudi *et al.*, *Virtual Phys. Prototyp.* **17**, 181–204 (2022).
16. S. Shevchik *et al.*, *Sci. Rep.* **10**, 3389 (2020).
17. Z. Smoqi *et al.*, *J. Mater. Process. Technol.* **304**, 117550 (2022).
18. S. A. Shevchik, C. Kenel, C. Leinenbach, K. Wasmer, *Addit. Manuf.* **21**, 598–604 (2018).
19. Materials and methods are available as supplementary materials.
20. L. Caprio, A. G. Demir, B. Previtali, *Addit. Manuf.* **36**, 101470 (2020).
21. B. J. Simonds *et al.*, *Phys. Rev. Appl.* **10**, 044061 (2018).
22. C. Zhao *et al.*, *Phys. Rev. X* **9**, 021052 (2019).
23. V. V. Semak, J. A. Hopkins, M. H. McCay, T. D. McCay, *J. Phys. D Appl. Phys.* **28**, 2443–2450 (1995).
24. N. Kouraytem *et al.*, *Phys. Rev. Appl.* **11**, 064054 (2019).
25. R. R. Selvaraju *et al.*, Grad-CAM: Visual explanations from deep networks via gradient-based localization, *2017 IEEE International Conference on Computer Vision (ICCV)* (2017), pp. 618–626.
26. T. G. Spears, S. A. Gold, *Integr. Mater. Manuf. Innov.* **5**, 16–40 (2016).
27. S. K. H. Everton, M. Hirsch, P. Stravroulakis, R. K. Leach, A. T. Clare, *Mater. Des.* **95**, 431–445 (2016).
28. M. Grasso, B. M. Colosimo, *Meas. Sci. Technol.* **28**, 044005 (2017).
29. B. Brandau, T. Mai, F. Brueckner, A. F. H. Kaplan, *Opt. Lasers Eng.* **155**, 107050 (2022).
30. M. Schürmann *et al.*, *Proc. SPIE* **9351**, 93510S (2015).
31. W. G. Guo, Q. Tian, S. Guo, Y. Guo, *CIRP Ann.* **69**, 205–208 (2020).

ACKNOWLEDGMENTS

We thank A. Deriy and V. Nikitin at the Advanced Photon Source for their technical support of the beamline experiments. We also thank B. Simonds at the National Institute of Standards and Technology, M. Bayat at the Technical University of Denmark, A. Mane and I. Muqica at Flow Science, and K. Jones at Carnegie Mello University for the fruitful discussions. This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science user facility operated for the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357. All data prepared, analyzed, and presented have been developed in a specific context of work and were prepared for internal evaluation and use pursuant to that work authorized under the referenced contract. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the US government, any agency thereof, or Honeywell Federal Manufacturing & Technologies. This publication has been authored by Honeywell Federal Manufacturing & Technologies under contract DE-NA0002839 with the DOE. A provisional application for a US patent (63427022) has been filed based on this work. **Funding:** This work was supported by the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357 and Honeywell Federal Manufacturing & Technologies under contract DE-NA0002839 with the DOE. **Author contributions:** Conceptualization: T.S. and Z.R. Methodology: T.S. and Z.R. Investigation: Z.R., T.S., L.G., S.J.C., K.F., P.S., A.D.R., L.C., and A.C. Visualization: Z.R., L.G., and T.S. Funding acquisition: W.E., A.C., T.S., and L.C. Project administration: T.S. and A.C. Supervision: T.S. Writing – original draft: Z.R. and T.S. Writing – review and editing: L.G., S.J.C., K.F., P.S., A.C., W.E., A.D.R., and L.C. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add4667
Materials and Methods
Supplementary Text
Figs. S1 to S26
Tables S1 to S7
References (*32*–*52*)
Movies S1 to S22

Submitted 13 June 2022; accepted 1 December 2022
10.1126/science.add4667

PLANT SCIENCE

Maize resistance to witchweed through changes in strigolactone biosynthesis

C. Li[1], L. Dong[1]*, J. Durairaj[2]†, J.-C. Guan[3], M. Yoshimura[4,5,6], P. Quinodoz[5], R. Horber[5], K. Gaus[5], J. Li[7], Y. B. Setotaw[7], J. Qi[7], H. De Groote[8], Y. Wang[1], B. Thiombiano[1], K. Floková[1,9], A. Walmsley[1], T. V. Charnikhova[1], A. Chojnacka[1], S. Correia de Lemos[2,10], Y. Ding[11], D. Skibbe[12], K. Hermann[5], C. Screpanti[5], A. De Mesmaeker[5], E. A. Schmelz[11], A. Menkir[13], M. Medema[2], A. D. J. Van Dijk[2], J. Wu[7], K. E. Koch[3], H. J. Bouwmeester[1]*

Maize (*Zea mays*) is a major staple crop in Africa, where its yield and the livelihood of millions are compromised by the parasitic witchweed *Striga*. Germination of *Striga* is induced by strigolactones exuded from maize roots into the rhizosphere. In a maize germplasm collection, we identified two strigolactones, zealactol and zealactonoic acid, which stimulate less *Striga* germination than the major maize strigolactone, zealactone. We then showed that a single cytochrome P450, ZmCYP706C37, catalyzes a series of oxidative steps in the maize-strigolactone biosynthetic pathway. Reduction in activity of this enzyme and two others involved in the pathway, ZmMAX1b and ZmCLAMT1, can change strigolactone composition and reduce *Striga* germination and infection. These results offer prospects for breeding *Striga*-resistant maize.

Food security is a growing challenge in the face of climate change and increasing food needs (*1*). Maize (*Zea mays*) is one of the most important staple crops in the world, especially in Africa. There, its yield is compromised by the parasitic witchweeds *Striga hermonthica* and *Striga asiatica*. Damage from these *Striga* species threatens the livelihood of millions of people, particularly in sub-Saharan regions (fig. S1) (*2*, *3*). *Striga* seeds lay dormant in soil until their germination is triggered by strigolactones (SLs), signaling compounds exuded by the roots of plants, including maize. The first known SL, strigol, was discovered in the 1960s in the root exudates of cotton (*4*). In addition to having been co-opted as a cue for root-parasitic plants, SLs serve as host signals for beneficial arbuscular mycorrhizal fungi (AMF) and are plant hormones with developmental roles (*5*–*9*).

Thus far, more than 35 different SLs have been discovered, all containing the conserved D-ring (Fig. 1A) (*10*–*12*). The canonical SLs include two groups, the "strigol-type" and "orobanchol-type," whereas noncanonical SLs lack the A-, B-, and/or C-rings (*10*–*12*). Plants usually exude a blend of different SLs, and the composition of the root exudate can vary greatly between and sometimes also within plant species. Many of the SLs display substantial differences in their biological activity, such as the induction of AMF hyphal branching and parasitic plant germination (*9*, *13*–*15*). The biological importance of SL blends is far from understood, but in sorghum (*Sorghum bicolor*), a change in SLs from 5-deoxystrigol to orobanchol decreased *Striga* germination and increased field resistance (*16*).

The mechanisms of SL biosynthesis have only been partially elucidated. Three enzymes—DWARF 27 (D27) and two carotenoid cleavage dioxygenases 1(CCDs), CCD7 and CCD8—catalyze the conversion of β-carotene to carlactone (CL) (Fig. 1A) (*17*, *18*). In *Arabidopsis*, CL is oxidized to form carlactonoic acid (CLA) by a cytochrome P450 (CYP) monooxygenase, CYP711A1, encoded by More Axillary Growth 1 (MAX1) homolog AtMAX1 (*19*). *Arabidopsis* has a single copy of this MAX1, whereas maize has three homologs, and rice has five (*18*, *20*). Although both the *Arabidopsis* AtMAX1 and the maize ZmMAX1b form CLA from CL, the rice MAX1 homologs, Os900 and Os1400, instead convert CL to 4-deoxyorobanchol (4DO) and orobanchol, respectively (*18*, *21*). Dicots also form orobanchol, but from CLA rather than CL, and with a different cytochrome P450, CYP722C. A homolog of this CYP722C can also produce 5-deoxystrigol from CLA (*22*, *23*).

[1]Plant Hormone Biology Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands. [2]Bioinformatics Group, Wageningen University & Research, 6708 PB Wageningen, Netherlands. [3]Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA. [4]Laboratorium für Organische Chemie, Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland. [5]Syngenta Crop Protection AG, Schaffhauserstrasse 101, CH-4332 Stein, Switzerland. [6]Kyoto University, iCeMS, Yoshida Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan. [7]Department of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China. [8]International Maize and Wheat Improvement Center (CIMMYT), PO Box 1041-00621, Nairobi, Kenya. [9]Laboratory of Growth Regulators, Institute of Experimental Botany, The Czech Academy of Sciences and Faculty of Science, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic. [10]Plant genomics and transcriptomics group, Institute of Biosciences, Sao Paulo State University, 13506-900 Rio Claro, Brazil. [11]Section of Cell and Developmental Biology, University of California at San Diego; La Jolla, CA 92093, USA. [12]Seeds Research, Syngenta Crop Protection, LLC, Research Triangle Park, NC 27709, USA. [13]International Institute of Tropical Agriculture, PMB 5320 Oyo Road, Ibadan, Nigeria.
***Corresponding author. Email: h.j.bouwmeester@uva.nl (H.J.B.), l.dong2@uva.nl (L.D.)**
†Present address: Biozentrum, University of Basel, Spitalstrasse 41, 4056 Basel, Switzerland.

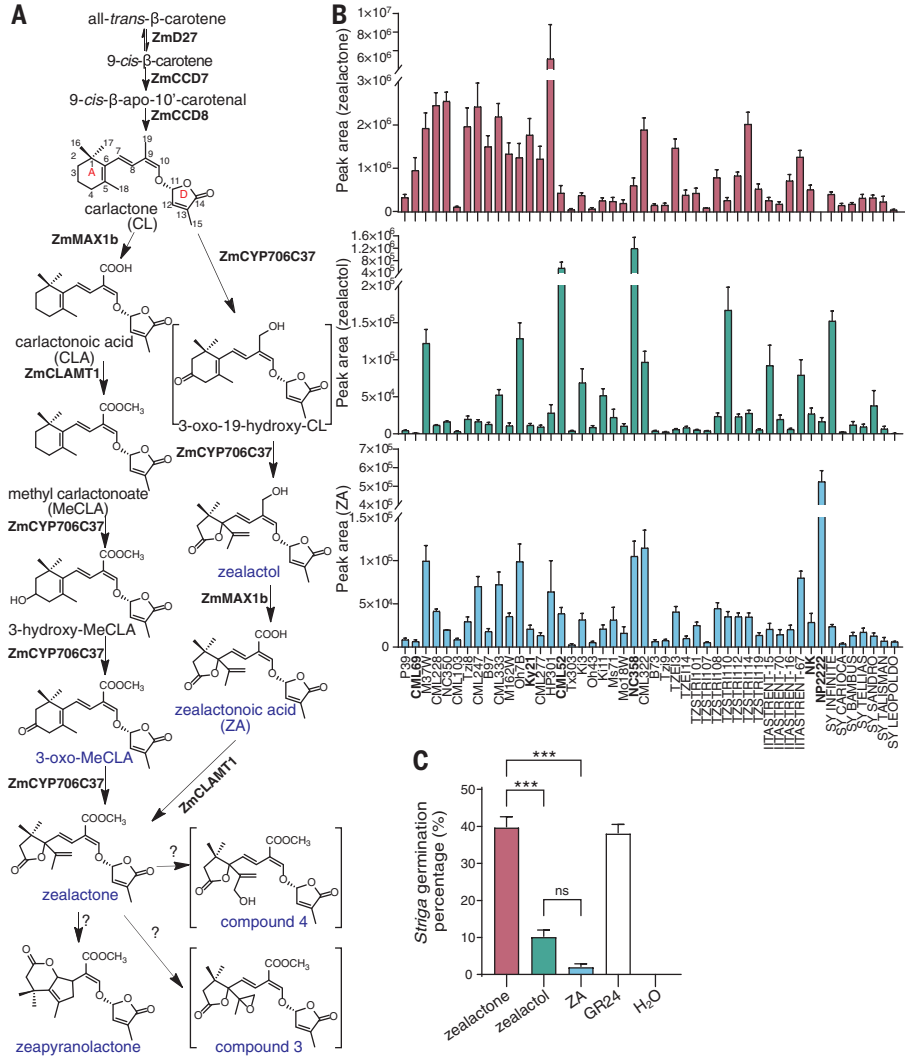


Fig. 1. Discovery of two strigolactones with low *Striga* germination–inducing activity from maize line screening. (**A**) Strigolactone (SL) biosynthetic pathway of maize. The enzymes identified in this study are shown in bold. SLs detected in maize root exudate are indicated in blue. Structures in square brackets are putative. (**B**) Detection of three maize SLs (zealactone, mass/charge ratio (*m/z*) 377 > 97; zealactol, *m/z* 331 > 97; ZA, *m/z* 363 > 249) in root exudate of a collection of maize lines. Names of lines selected for further analysis are indicated in bold. Data for the other four maize SLs are shown in fig. S3. (**C**) Induction of germination of *Striga* by zealactone, zealactol, and ZA (0.347 μM). GR24 (0.335 μM) and water were used as positive and negative control, respectively. Bars indicate means ± SEM. Ns, not significant ($P > 0.05$), ***$P < 0.001$, one-way ANOVA test followed by Tukey's multiple comparisons test comparing the mean of each column with the mean of every other column.

Maize roots exude at least six SLs, two of which have been structurally identified as zealactone and zeapyranolactone (Fig. 1A) (*24*–*26*). However, the identities of the other four SLs remained elusive, as well as the biosynthetic differences between the six and their individual roles in *Striga* germination. In this study, we reveal natural variation in the maize SL blend, identify three new maize SLs, elucidate the entire maize SL biosynthetic pathway, and show that changes in the composition of the SL blend correspond to differences in *Striga* germination and infection. These findings create a pathway for reducing the notorious agricultural problem of *Striga* infection through breeding maize for favorable SL composition.

Natural variation in strigolactone production by maize

To assess the extent of variation in the production of SLs by maize, we grew a collection of maize genotypes, sampled their root exudate, and analyzed SLs with multiple reaction monitoring (MRM) liquid chromatography–tandem mass spectroscopy (LC/MS/MS) (Fig. 1B and figs. S2 and S3) (*24*, *25*). Quantities of exuded SLs varied among these lines (Fig. 1B and fig. S3). Moreover, one of the genotypes, NP2222, displayed a distinctive SL profile, lacking detectable levels of all but two SLs, an unknown SL and designated compound 5 (Fig. 1B and fig. S3). Compound 5 was previously noted in maize root exudate (*24*), but its low abundance and chemical instability hampered structural characterization. Therefore, on the basis of nuclear magnetic resonance (NMR) spectra and retrosynthetic analysis (*24*, *27*–*29*), we postulated structures and subsequently synthesized compound 5 as well as the other unknown SL (figs. S4 to S12). The synthetic products were identical to the natural ones in maize root exudate and were designated zealactol (compound 5) and zealactonoic acid (ZA) (the other unknown SL) (figs. S9 and S12). Bioassay of *Striga* germination showed that both zealactol and ZA were less inductive than zealactone (Fig. 1C), an outcome that highlights how strongly minute differences in SL structure can alter their biological activity. These findings are further supported by work on sorghum (*16*). To unravel the mechanistic basis for these differences in SL blends, we revealed the biosynthetic pathway of maize SLs.

Three maize genes encode the carlactone biosynthetic pathway

Through homology, we identified the maize orthologs *D27*, *CCD7*, and *CCD8*, which catalyze the formation of CL from β-carotene in other plant species (tables S1 and S2). To confirm *ZmCCD8* function, we analyzed root exudate of two independent *zmccd8* mutants (in W22 and Mo17 backgrounds) (*30*). Zealactone was not detected, although it was the major SL in wild-type exudate (fig. S13A), showing that *ZmCCD8* is a key enzyme in maize SL biosynthesis (*17*, *31*, *32*). The transient expression of *ZmD27* (*GRMZM2G158175*), *ZmCCD7* (*GRMZM2G158657*), and *ZmCCD8* (*GRMZM2G446858*) together in *Nicotiana benthamiana* led to accumulation of CL (Figs. 1A and 2A, fig. S14A, and table S3), which is consistent with results from rice and tomato orthologs (*21*, *33*).

Identification of gene candidates for carlactone conversion

On the basis of the structures of the maize SLs identified thus far (Fig. 1A and fig. S2) (*24*–*26*), we postulated the involvement of a methyl transferase and several CYPs in the pathway downstream of CL. Several bioinformatic approaches were combined to select candidate genes for further functional characterization.

Mutual Rank (MR)–based global gene coexpression analysis (*34*, *35*) showed that of the three maize *MAX1* homologs, only *ZmMAX1b* tightly coexpressed with *ZmCCD8* (fig. S15), making it the strongest candidate for the next

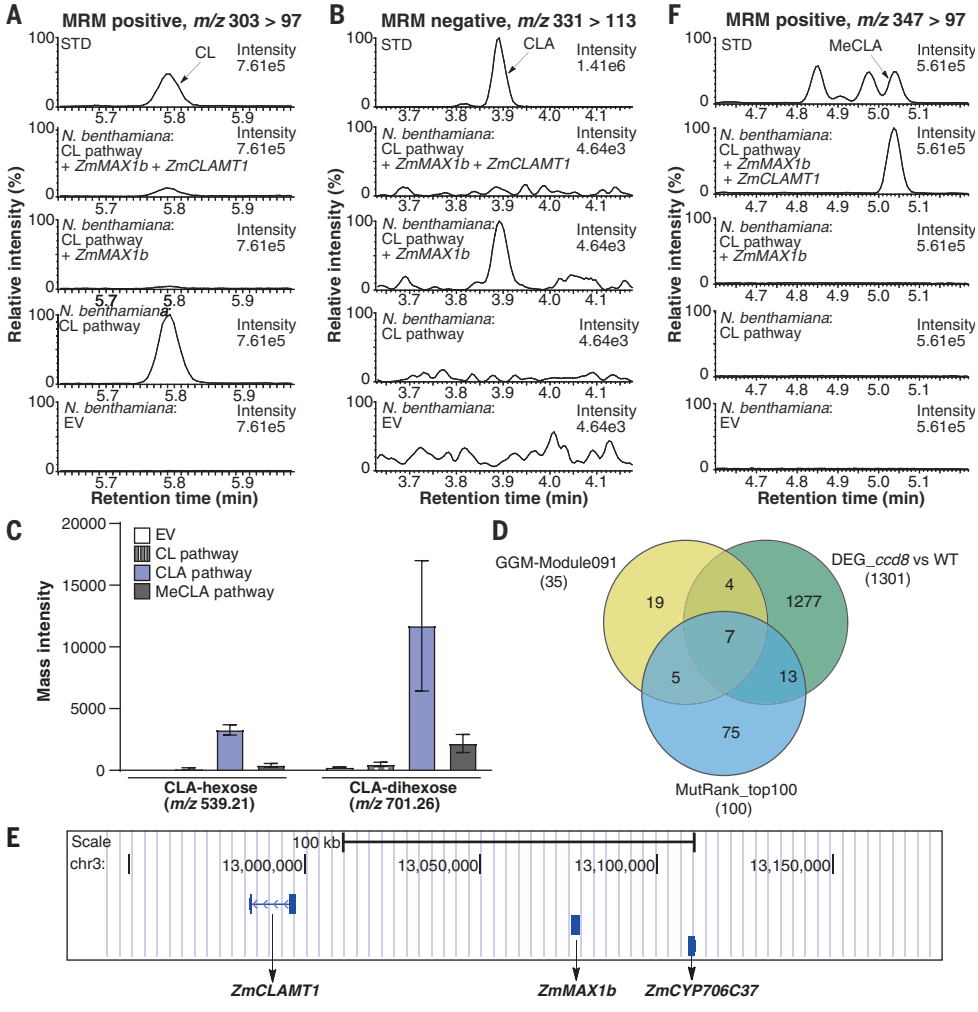


Fig. 2. Identification of gene candidates for maize strigolactone biosynthesis. (**A** and **B**) Representative MRM-LC/MS/MS chromatograms of carlactone (CL), $[M+H]^+$ *m/z* 303 > 97 (A), and carlactonoic acid (CLA), $[M-H]^-$ *m/z* 331 > 113 (B), in *N. benthamiana* leaf samples transiently expressing maize strigolactone (SL) precursor pathway genes. (**C**) Untargeted metabolomics to identify CLA conjugates in *N. benthamiana* leaf samples. *m/z* 539.21: CLA + hexose + formic acid − H_2O; *m/z* 701.26: CLA + 2 hexose + formic acid − H_2O (**D**) Venn diagram of candidate gene numbers from several analyses: module091 from maizeGGM, genes differentially expressed in *zmccd8* roots (compared with wild type), and the top 100 genes coexpressed with *ZmCCD8* and *ZmMAX1b* (*34*, *35*). (**E**) Putative SL biosynthetic gene cluster on chromosome 3 consisting of *ZmCLAMT1*, *ZmMAX1b*, and *ZmCYP706C37*, adapted from screenshot from UCSC Genome Browser on *Z. mays* (B73 RefGen_v3) Assembly (zm3) (http://genome.ucsc.edu) (*39*). (**F**) Representative chromatograms of methylcarlactonoic acid (MeCLA), $[M+H]^+$ *m/z* 347 > 97, in *N. benthamiana* leaf samples. STD, standard; EV, empty vector infiltrated control sample. CL pathway, maize carlactone biosynthetic pathway genes, *ZmD27*, *ZmCCD7*, and *ZmCCD8*. CLA pathway, CL pathway genes + *ZmMAX1b*. MeCLA pathway, CLA pathway genes + *ZmCLAMT1*. Bars indicate mean ± SEM.

biosynthetic step. Analysis of root exudate from a *zmmax1a zmmax1c* double mutant (supplementary materials) showed wild-type levels of zealactone, thus excluding both homologs from being the biosynthetic genes we sought (fig. S13B). Earlier research also demonstrated that ZmMAX1b (GRMZM2G023952) converts CL to CLA more efficiently than does ZmMAX1a (GRMZM2G018612) or ZmMAX1c (GRMZM2G070508) (*18*). The amounts of CL in leaf extracts decreased after coinfiltration of *ZmMAX1b* with *ZmD27*, *ZmCCD7*, and *ZmCCD8* in *N. benthamiana,* (Fig. 2A), confirming that ZmMAX1b uses CL as a substrate (*18*). However, only traces of the expected product, CLA, were detected in this expression system (Fig. 2B and fig. S14B). To resolve this enigma, *N. benthamiana* extracts were analyzed with LC–quadrupole time-of-flight (QTOF)–MS. Prominent peaks of CLA-hexose and CLA-dihexose conjugates were detected in samples expressing the maize CL pathway genes together with *ZmMAX1b*. These conjugates were lacking in control samples and other gene combinations (Fig. 2C and table S4). Similar conjugation has been demonstrated for the transient production of other acidic compounds with *N. benthamiana* (*36*, *37*).

For selection of remaining candidate genes, we combined three approaches: (i) MR–based coexpression with *ZmCCD8* and *ZmMAX1b* as baits (fig. S15), (ii) coexpression modules in MaizeGGM2016 (*38*), and (iii) differential gene expression in a *zmccd8* mutant (Fig. 2D). For the latter, we assumed that SL pathway genes downstream of *CCD8* would be transcriptionally regulated in the *zmccd8* mutant (*33*). The *ZmCCD7*, *ZmCCD8*, and *ZmMAX1b* genes clustered together in MaizeGGM2016 module 091, suggesting that the 32 other genes in this module were candidates for the missing pathway genes (table S5). In the roots of *zmccd8* seedlings, 1301 genes were differentially expressed (DEGs) (less than or equal to twofold change, false discovery rate (FDR) < 0.05) compared with the B73 wild type (tables S5 and S6). These three approaches shared a seven-gene overlap (Fig. 2D and table S2) in which three [*GRMZM2G033126*, *GRMZM2G158342*, and *GRMZM2G023952* (*ZmMAX1b*)] formed a putative gene cluster on chromosome 3 (Fig. 2, D and E, and fig. S15) (*39*). Genes homologous to these also cluster in other *Poaceae* species (fig. S16), but the functional importance is unknown. So too is the identity of SLs produced by some of these species, such as switchgrass.

ZmCLAMT1 is a carlactonoic acid methyltransferase

Because SLs zealactone and zeapyranolactone are methylesters, their proposed precursor has been methyl carlactonoate (MeCLA) (*24*). Thus, we sought a methyltransferase gene that causes the formation of MeCLA from CLA. We bioinformatically identified a top candidate (*GRMZM2G033126*) (Fig. 2, D and E), which successfully produced MeCLA in *N. benthamiana* when transiently expressed together with genes for the maize CLA pathway (Fig. 2F). We therefore identified *GRMZM2G033126* as a carlactonoic acid methyltransferase gene and named the enzyme ZmCLAMT1 (Fig. 1A). The maize gene is an ortholog of *At4g36470*, which was recently found to encode a carlactonoic acid methyltransferase CLAMT in *Arabidopsis* (*40*, *41*).

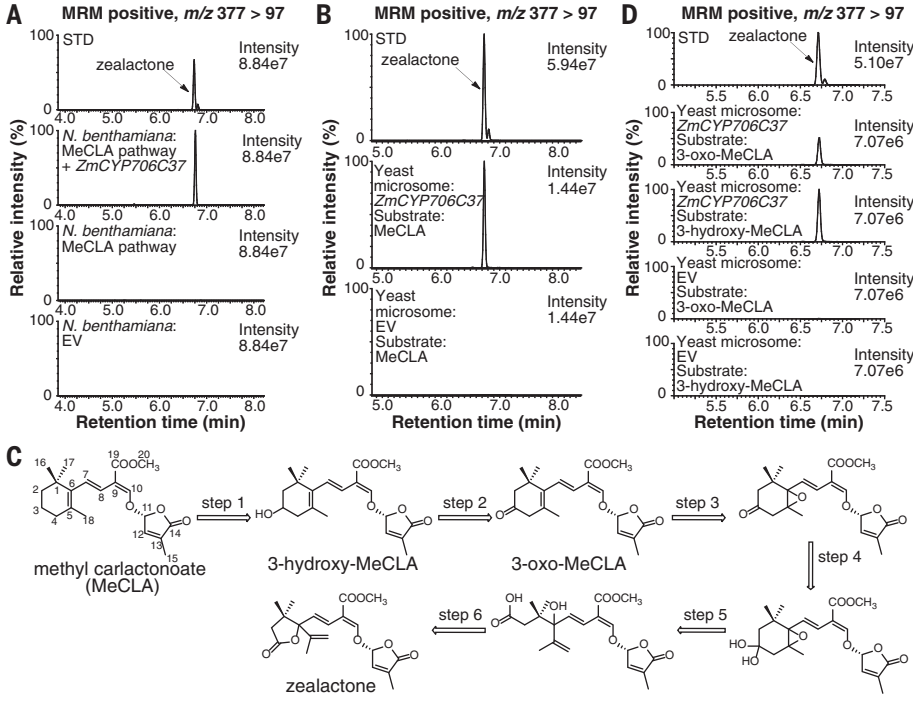


Fig. 3. Zealactone biosynthesis. (**A**) Representative MRM–LC/MS/MS chromatograms of zealactone, $[M+H]^+$*m/z* 377 > 97, in *N. benthamiana* leaf samples. (**B** and **D**) Representative MRM–LC/MS/MS chromatograms of zealactone from in vitro assays with yeast microsomes expressing *ZmCYP706C37* or empty vector (EV) with methyl carlactonoate (MeCLA), 3-hydroxy-MeCLA, or 3-oxo-MeCLA as substrate. (**C**) Proposed enzymatic conversion of methyl carlactonoate (MeCLA) to zealactone.

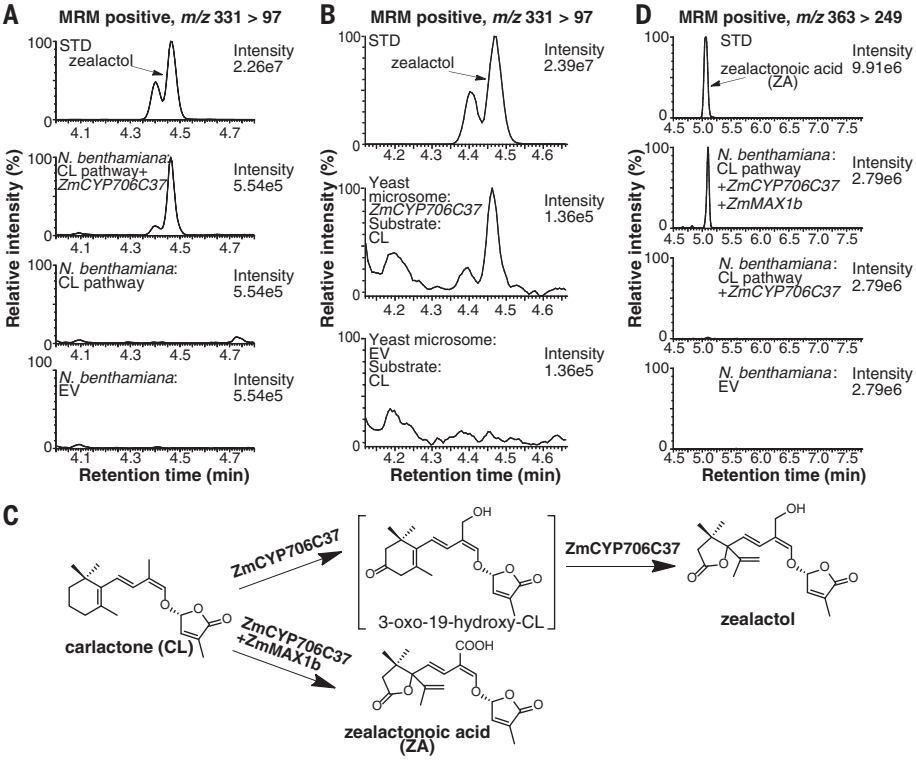


Fig. 4. Zealactol and zealactonoic acid biosynthesis. (**A**) Representative MRM–LC/MS/MS chromatograms of zealactol, $[M+H-H_2O]^+$*m/z* 331 > 97, in *N. benthamiana* leaf samples. (**B**) Representative MRM–LC/MS/MS chromatograms of zealactol from in vitro assays with yeast microsomes expressing *ZmCYP706C37* or empty vector (EV) with carlactone (CL) as substrate. (**C**) Reactions from CL to zealactol and ZA catalyzed by ZmCYP706C37 and ZmMAX1b. Structure in square brackets is putative. (**D**) Representative MRM–LC/MS/MS chromatograms of ZA, $[M+H]^+$*m/z* 363 > 249, in *N. benthamiana* leaf samples. STD, standard; EV, empty vector control. CL pathway, maize carlactone biosynthetic pathway genes, *ZmD27*, *ZmCCD7*, and *ZmCCD8*.

ZmCYP706C37 catalyzes formation of several maize strigolactones

The other candidate genes were coinfiltrated by different combinations of precursor-pathway genes. Coinfiltration of *ZmCYP706C37* (*GRMZM2G158342*) (*42*) by those encoding the MeCLA pathway decreased levels of MeCLA, indicating that this CYP can use MeCLA as a substrate (fig. S17A) and produce zealactone (Fig. 3A and fig. S2). To check for other possible biosynthetic pathways, we also coexpressed *ZmCYP706C37* with genes encoding the CL pathway enzymes. This combination resulted in production of zealactol (Fig. 4A and fig. S17B). Formation of both zealactone and zealactol involves complex rearrangement of the SL A ring and, for zealactol, a hydroxylation at C19 as well. To exclude the possibility of endogenous enzymes from *N. benthamiana* contributing to these complex conversions, we expressed *ZmCYP706C37* in yeast, isolated its microsomes, and analyzed product formation with different substrates (Figs. 3B and 4B). This approach confirmed that ZmCYP706C37 can

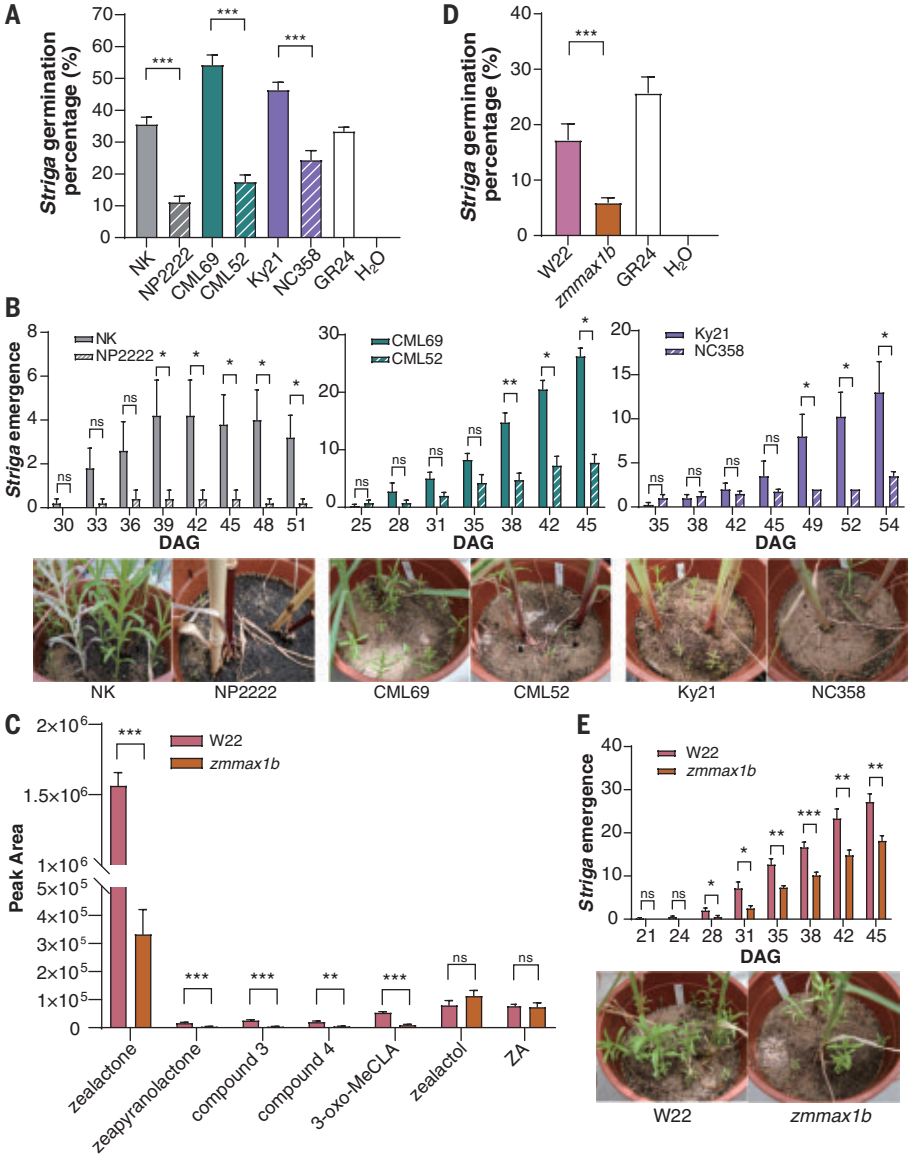


Fig. 5. Changes in the maize strigolactone blend result in changes in *Striga* resistance. (**A** and **D**) Induction of *Striga* germination by root exudates of selected maize lines. GR24 (0.335 μM) and water were used as positive and negative control, respectively. (**B** and **E**) *Striga* infection of selected maize lines. Emerged *Striga* numbers were recorded; representative photos highlight the differences. DAG, days after germination of maize. (**C**) SL levels in the root exudate of *zmmax1b* and its wild type, W22. Bars indicate means ± SEM, ns = not significant ($P > 0.05$), *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, two-tailed, unpaired *t* test.

convert MeCLA to zealactone and CL to zealactol (Fig. 1A).

To form zealactone from MeCLA, ZmCYP706C37 must catalyze several consecutive oxidative reactions with 3-hydroxy-MeCLA and 3-oxo-MeCLA as putative intermediates (Figs. 1A and 3C). The latter two compounds were previously synthesized as intermediates in the total synthesis of heliolactone (*43*). We used them here as substrates in our *ZmCYP706C37*-expressing yeast-microsome assay, and both were successfully converted to zealactone (Fig. 3D). We developed an MRM method for detection of these compounds (fig. S2) and identified them as intermediate products in the conversion of MeCLA to zealactone (fig. S18). Moreover, analysis of maize root exudate revealed that 3-oxo-MeCLA is also a natural maize SL previously referred to as compound 6 (fig. S19 and Fig. 1A) (*24*). These results demonstrate that a single enzyme, ZmCYP706C37, can catalyze the many oxidative steps necessary for the conversion of MeCLA to zealactone that were previously hypothesized to require several enzymes (Figs. 1A and 3C) (*24*).

For additional insight into the parallel biosynthetic pathway of CL to zealactol, we further analyzed samples from *N. benthamiana* and yeast microsome assays with untargeted metabolomics and MRM-LC-MS/MS. This process revealed another putative intermediate, 3-oxo-19-hydroxy-CL (compound 7) (Fig. 1A and figs. S2 and S20 and table S7). LC-QTOF-MS analysis showed that the accurate mass of compound 7 is consistent with its putative structure (fig. S20). On the basis of these data, we included compound 7 as an intermediate in the postulated steps required to convert CL to zealactol (Fig. 4C and fig. S21). Moreover, agroinfiltration of the CL pathway genes with *ZmCYP706C37* and *ZmMAX1b* resulted in production of ZA, a result also confirmed with LC-QTOF-MS (Fig. 4, C and D, and fig. S22).

Last, analysis of root exudate from a *zmcyp706c37* mutant [EMS4-045ad8, stop-codon gained (fig. S23A)] showed no detectable levels of zealactol, ZA, zealactone, or three other SLs derived from the latter (fig. S23B) (*44*). Although 3-oxo-MeCLA was detectable in the mutant exudate, it was present at a much lower level than in that of the wild type. Instead, CLA and MeCLA accumulated in the mutant exudate, whereas they are absent in the wild type exudate (fig. S23, C and D). Together, these data support our functional characterization of ZmCYP706C37.

Biosynthetic control of the maize strigolactone blend

To determine how the different maize SLs are biosynthetically related, we applied 3-hydroxy-MeCLA, 3-oxo-MeCLA, and zealactol to seedlings of another commercial line, NK Falkone, which were treated with fluridone, an inhibitor of SL biosynthesis (*45*). Each of these three compounds complemented zealactone production (fig. S24A), confirming that they can serve as biosynthetic precursors for zealactone. Combined transient expression of *ZmMAX1b* and *ZmCLAMT1* in *N. benthamiana* leaves and subsequent infiltration of zealactol also showed that the latter can be converted to zealactone by ZmMAX1b together with ZmCLAMT1 (Fig. 1A and fig. S25). Application of zealactone to fluridone-treated plants led to the formation of zeapyranolactone and two other maize SLs, designated compounds 3 and 4, suggesting that zealactone is their precursor (Fig. 1A and fig. S24, B to D) (*24*).

Next, we sought mechanisms underlying the distinctive maize SL profile of NP2222 (fig. S26). This line produces zealactone in fluridone-treated seedlings, as does NK Falkone, but only from MeCLA and 3-oxo-MeCLA, not from zealactol (figs. S24A and S26A), suggesting inactivity of MAX1b and/or CLAMT1. As previously noted, ZA accumulated in the root exudate of NP2222 (Fig. 1B and fig. S26D), indicating dysfunction of CLAMT1. Zealactol

added to either NK Falkone or NP2222 was converted to ZA, showing that ZmMAX1b is active in NP2222 (fig. S26, B and C). Inspection of the *CLAMT1* sequence in a proprietary NP2222 genome database revealed a large insertion in the second exon of this gene, and reverse transcriptase polymerase chain reaction (RT-PCR) showed that regions flanking the insertion were not transcribed (fig. S26E). These collective data indicate disfunction of CLAMT1 in NP2222.

To analyze biological consequences of the different SL profiles, several maize lines were selected for *Striga* germination and infection assays. The NP2222 root exudate induced much lower germination than that of NK Falkone. Results were consistent with their respective SL profiles and differences in germination-inducing activity of the individual SLs (Figs. 1C and 5A and fig. S26D). CML52 and NC358, both with high proportions of zealactol and ZA, induced significantly less *Striga* germination than did CML69 and Ky21, which produced mostly zealactone despite similar total SL peak areas (Figs. 1C and 5A, and fig. S27, A and B). These differences were also reflected in a *Striga* infection assay with a containerized system, in which *Striga* emergence was less for low-zealactone genotypes (Fig. 5B). In addition to their SL blend, these lines may have other genetic differences that could affect these results. However, we also analyzed a gene-suppression mutant of *ZmMAX1b* (transposon insertion in a W22 background) (fig S28, A and B). This mutant exuded significantly less izealactone and zealactone-derived SLs, whereas the level of zealactol was higher than in the W22 control (Fig. 5C). The *zmmax1b* mutant also induced less *Striga* germination and emergence (Fig. 5E). Results confirm that a change in activity of specific SL biosynthetic enzymes in maize can change the SL composition and confer *Striga* resistance. Although the underlying mechanisms are completely different, these findings resemble those of *lgs* sorghum (*16*) and present a promising prospect for *Striga* resistance breeding in maize. The *zmmax1b* mutant did not exhibit a branching phenotype, in contrast to *zmccd8* (fig. S28C). Also, *zmcyp706c37*, which is located parallel to or downstream of ZmMAX1b, did not display an obvious branching phenotype either. This all suggests that the downstream SLs are not nor precursors of the branching inhibiting hormone and are therefore safe breeding targets that will not result in unwanted pleiotropic effects.

Conclusions

We have shown that two parallel SL biosynthetic pathways operate in maize and that both pathways produce the major maize SL, zealactone. Changes in flux through these pathways can alter the maize SL profile by shifting the balance between zealactone and zealactol plus ZA. Zealactol and ZA induce much less *Striga* germination, thus imparting a strong reduction in *Striga* infection to genotypes that exude more zealactol and ZA than zealactone. Future research should investigate whether these changes in the SL blend affect colonization by AM fungi, which was not observed for *lgs* sorghum (*16*). Our results offer a perspective for breeding *Striga* resistance through modification of the SL blend in maize and thus potentially reducing the devastating effects of this parasitic weed in Africa.

REFERENCES AND NOTES

1. T. Wheeler, J. von Braun, *Science* **341**, 508–513 (2013).
2. B. Badu-Apraku, F. M.A.B., *Advances in Genetic Enhancement of Early and Extra-Early Maize for Sub-Saharan Africa* (Springer Cham, 2017).
3. I. Dörr, *Ann. Bot. (Lond.)* **79**, 463–472 (1997).
4. C. E. Cook, L. P. Whichard, B. Turner, M. E. Wall, G. H. Egley, *Science* **154**, 1189–1190 (1966).
5. A. Besserer *et al*., *PLOS Biol.* **4**, e226 (2006).
6. V. Gomez-Roldan *et al*., *Nature* **455**, 189–194 (2008).
7. S. Al-Babili, H. J. Bouwmeester, *Annu. Rev. Plant Biol.* **66**, 161–186 (2015).
8. M. Umehara *et al*., *Nature* **455**, 195–200 (2008).
9. K. Akiyama, K. Matsuzaki, H. Hayashi, *Nature* **435**, 824–827 (2005).
10. H. Bouwmeester, C. Li, B. Thiombiano, M. Rahimi, L. Dong, *Plant Physiol.* **185**, 1292–1308 (2021).
11. K. Yoneyama *et al*., *J. Exp. Bot.* **69**, 2231–2239 (2018).
12. K. Mashiguchi, Y. Seto, S. Yamaguchi, *Plant J.* **105**, 335–350 (2021).
13. K. Akiyama, S. Ogasawara, S. Ito, H. Hayashi, *Plant Cell Physiol.* **51**, 1104–1117 (2010).
14. N. Mori, K. Nishiuma, T. Sugiyama, H. Hayashi, K. Akiyama, *Phytochemistry* **130**, 90–98 (2016).
15. H. I. Kim *et al*., *J. Pestic. Sci.* **35**, 344–347 (2010).
16. D. Gobena *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4471–4476 (2017).
17. A. Alder *et al*., *Science* **335**, 1348–1351 (2012).
18. K. Yoneyama *et al*., *New Phytol.* **218**, 1522–1533 (2018).
19. S. Abe *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18084–18089 (2014).
20. C. Cardoso *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2379–2384 (2014).
21. Y. Zhang *et al*., *Nat. Chem. Biol.* **10**, 1028–1033 (2014).
22. T. Wakabayashi *et al*., *Sci. Adv.* **5**, eaax9067 (2019).
23. T. Wakabayashi *et al*., *Planta* **251**, 97 (2020).
24. T. V. Charnikhova *et al*., *Phytochemistry* **137**, 123–131 (2017).
25. T. V. Charnikhova *et al*., *Phytochem. Lett.* **24**, 172–178 (2018).
26. X. Xie *et al*., *J. Pestic. Sci.* **42**, 58–61 (2017).
27. M. Yoshimura *et al*., *Helv. Chim. Acta* **103**, e2000017 (2020).
28. M. C. Dieckmann, P.-Y. Dakas, A. De Mesmaeker, *J. Org. Chem.* **83**, 125–135 (2018).
29. T. Kumagai *et al*., *Heterocycles* **36**, 1729–1734 (1993).
30. J. C. Guan *et al*., *Plant Physiol.* **160**, 1303–1317 (2012).
31. K. C. Snowden *et al*., *Plant Cell* **17**, 746–759 (2005).
32. W. Kohlen *et al*., *New Phytol.* **196**, 535–547 (2012).
33. Y. Zhang *et al*., *New Phytol.* **219**, 297–309 (2018).
34. E. Poretsky, A. Huffaker, *PeerJ* **8**, e10264 (2020).
35. S. Stelpflug *et al*, *The Plant Genome* **9**, plantgenome2015.04.0025 (2016).
36. L. Dong *et al*., *Metab. Eng.* **20**, 198–211 (2013).
37. X. Xu *et al*., *J. Exp. Bot.* **72**, 5462–5477 (2021).
38. S. Ma, Z. Ding, P. Li, *BMC Plant Biol.* **17**, 131 (2017).
39. W. J. Kent *et al*., *Genome Res.* **12**, 996–1006 (2002).
40. T. Wakabayashi *et al*., *Planta* **254**, 88 (2021).
41. K. Mashiguchi *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2111565119 (2022).
42. Y. Li, K. Wei, *BMC Plant Biol.* **20**, 93 (2020).
43. M. Yoshimura *et al*., *Helv. Chim. Acta* **102**, e1900211 (2019).
44. X. Lu *et al*., *Mol. Plant* **11**, 496–504 (2018).
45. J. A. López-Ráez *et al*., *New Phytol.* **178**, 863–874 (2008).

ACKNOWLEDGMENTS

We acknowledge S. Al Babili from King Abdullah University of Science and Technology and D. Werck-Reichhart from the University of Strasbourg for helpful discussions, as well as L. Hagmann from Syngenta for his support in NMR analyses and interpretation. **Funding:** This work was funded by the China Scholarship Council (CSC) PhD scholarship 201706300041 (C.L.), the European Research Council (ERC) Advanced grant CHEMCOMRHIZO 670211 (H.J.B.), the Dutch Research Council (NWO/OCW) Gravitation program Harnessing the second genome of plants (MiCRop) 024.004.014 (H.J.B.), the Marie Curie fellowship NEMHATCH 793795 (L.D.). K.E.K and J.G. acknowledge funding from the US National Science Foundation (NSF) Plant Genome Research Program (PGRP) (1421100 and 1748105). **Author contributions:** C.L., L.D., and H.J.B. conceived and designed the project. C.L. discovered and characterized the candidate genes, grew the plants, collected and analyzed the root exudate, cloned the genes, performed agroinfiltration, yeast microsome assay and plant compound treatment assays, and coordinated the project; K.F., T.V.C. and A.C. developed LC-MS methods and helped with SL analysis; T.V.C., J.D., and A.D.J.V.D. helped to establish the biosynthesis mechanisms; J.G. and K.E.K. developed and provided maize seeds (NAM, *zmccd8*, *zmmax1azmmax1c*, and *zmmax1b*) and analyzed RNA-seq and related data. B.T. and L.D. supported the metabolomics analysis; M.Y., K.G., A.D.M. synthesized zealactol and provided zealactone, 3-hydroxy-MeCLA, and 3-oxo-MeCLA; P.Q., R.H., and A.D.M. synthesized zealactonoic acid; J.L., Y.B.S., J.Q., and J.W. grew the *zmcyp706c37* EMS mutants, performed genotyping, selfing, and root exudate collection; H.D.G. collected and prepared the maps of maize and *Striga* occurrence. Y.W. helped with the agroinfiltration and yeast microsome assays; C.L., A.W., and B.T. performed the *Striga* germination and infection bioassays; S.M.C.d.L. and M.H.M. carried out the gene cluster analysis; Y.D. and E.A.S. provided support on coexpression analysis; D.K., K.H. and C.S. provided all commercial maize seeds from Syngenta and coordinated the collaboration with Syngenta. A.M. provided African inbred maize lines. C.L., L.D., and H.J.B. wrote the manuscript, with contributions from other authors. **Competing interests:** M.H.M. is a consultant to Corteva Agriscience, but that company was not involved in this work. All the other authors declare that they have no competing interests. **Data and materials availability:** The maize mutants *zmccd8* and *zmmax1azmmax1c* were obtained via a material transfer agreement (MTA) with the University of Florida Board of Trustees. The RNA-seq data of *zmccd8* and B73 root tissues are available in the NCBI database (BioProject PRJNA757767) under accession numbers of SRR15613590, SRR15613591, SRR15613599, SRR15613593, SRR15613594, and SRR15613595. All the other data are presented in the main text and in the Supplementary Materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq4775
Materials and Methods
Figs. S1 to S28
Tables S1 to S8
References (*46*–*70*)

Submitted 12 April 2022; accepted 30 November 2022
10.1126/science.abq4775

MPOX

Structure of monkeypox virus DNA polymerase holoenzyme

Qi Peng[1]†, Yufeng Xie[2]†, Lu Kuai[1]†, Han Wang[1,3], Jianxun Qi[1], George F. Gao[1,2,4,5,6]*, Yi Shi[1,5,6]*

The World Health Organization declared mpox (or monkeypox) a public health emergency of international concern in July 2022, and prophylactic and therapeutic measures are in urgent need. The monkeypox virus (MPXV) has its own DNA polymerase F8, together with the processive cofactors A22 and E4, constituting the polymerase holoenzyme for genome replication. Here, we determined the holoenzyme structure in complex with DNA using cryo–electron microscopy at the global resolution of ~2.8 angstroms. The holoenzyme possesses an architecture that suggests a "forward sliding clamp" processivity mechanism for viral DNA replication. MPXV polymerase has a DNA binding mode similar to that of other B-family DNA polymerases from different species. These findings reveal the mechanism of the MPXV genome replication and may guide the development of anti-poxvirus drugs.

As of 2 December 2022, over 82,000 human mpox (or monkeypox) cases have been laboratory confirmed in 110 countries worldwide (https://www.cdc.gov/). Most infection cases have been reported in Europe and other non-endemic countries, including China (*1*), and these cases were mostly found in homosexual young men (*2*). Human-to-human transmission usually occurs through close contact with lesions, respiratory droplets, body fluids, and contaminated materials, such as bedding (*3*).

Although the monkeypox virus (MPXV) was first isolated from a monkey in Denmark in 1958, its natural host was thought to be rodent (*3*, *4*). Since the first human mpox case was identified in the Democratic Republic of the Congo (*5*), it has been endemic to several central and western African countries (*6*, *7*). Sporadic infection cases have been reported outside Africa, including England, the United States, Singapore, and Israel, and are mainly associated with travelers from endemic countries, nosocomial infections, or direct contact with imported rodents infected with MPXV (*4*, *8*, *9*). Phylogenetic analysis has revealed that MPXV can be classified into two genetic clades: the West Africa clade and the more pathogenic Congo Basin clade (*10*, *11*). The 2022 outbreak of MPXV belongs to the West Africa clade and most likely has a single origin that has not been identified (*12*).

[1]CAS Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. [2]Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China. [3]College of Future Technology, Peking University, Beijing 100871, China. [4]Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China. [5]Center for Influenza Research and Early-warning (CASCIRE), CAS-TWAS Center of Excellence for Emerging Infectious Disease (CEEID), Chinese Academy of Sciences, Beijing 100101, China. [6]Research Unit of Adaptive Evolution and Control of Emerging Viruses, Chinese Academy of Medical Sciences, Beijing 100052, China.
***Corresponding author. Email: gaof@im.ac.cn (G.F.G.); shiyi@im.ac.cn (Y.S.)**
†These authors contributed equally to this work.

MPXV is a large double-stranded DNA virus that replicates exclusively in the cytoplasm of the infected cells. It belongs to the *Orthopoxvirus* genus of the Poxviridae family, which also includes the variola virus that causes smallpox and has killed millions of humans in recorded history. Similar to the vaccinia virus (VACV), the prototype of poxviruses, MPXV may enter host cells by either fusion with the plasma membrane or endocytosis, and at least 16 proteins in the virion membrane are involved in the entry process (*13*). After entry, the virus initiates early gene transcription events, and viral DNA synthesis occurs at perinuclear sites called viral factories (*14*, *15*). The MPXV replicative holoenzyme consists of catalytic polymerase F8 (equivalent to E9 in VACV), a heterodimeric processivity factor consisting of A22 (equivalent to A20 in VACV) and uracil-DNA glycosylase E4 (equivalent to D4 in VACV).

Previous genetic, biochemical, and structural studies on the VACV E9-A20-D4 core replication machinery have advanced our understanding of poxvirus DNA replication. VACV E9 was recognized as a member of the B-family DNA polymerase, and structural analysis has revealed the canonical features of DNA polymerases and five poxvirus-specific insertions (*16*). The E9 polymerase alone does not have processive DNA synthesis activity unless it is bound to its heterodimeric cofactor A20/D4 (*17*–*20*). Although poxvirus DNA polymerase shares many features with other B-family polymerases, the processivity factor is distinctive. In VACV, A20 serves as an essential bridge to link E9 and D4 together and shares no homology with viral proteins beyond poxvirus. The N-terminal domain of A20 binds to D4 (*21*–*23*), and its C-terminal domain binds to one insertion in the palm domain of E9 (*24*). Given that the DNA replication machinery is extremely conserved for orthopoxviruses, with a sequence identity of more than 97% between VACV and MPXV, the results obtained for VACV could also be applied to MPXV. However, we are still awaiting a reliable high-resolution structure of the replicating state of the *Orthopoxvirus* polymerase holoenzyme, and the mode of operation of the processivity factor needs to be elucidated.

Results

Biochemical characterization of the purified polymerase proteins

We coexpressed MPXV F8 polymerase and the A22-E4 heterodimer using the baculovirus expression system and purified the homogeneous F8-A22-E4 heterotrimer protein for enzymatic and structural studies (fig. S1). When a 38-nucleotide (nt) template DNA was used with a 24-nt primer DNA, the wild-type holoenzyme heterotrimer displayed weak primer extension activity in the reaction buffer with deoxynucleotide triphosphate (dNTP) (fig. S1). Moreover, exonuclease activity was confirmed using an enzymatic assay without dNTP substrates. The holoenzyme could completely degrade the primer DNA in an adenosine 5′-triphosphate (ATP)–independent manner (fig. S1). We also prepared an exonuclease-deficient F8 mutant protein and found that the F8-mutant-A22-E4 holoenzyme did not cleave the primer-template DNA, thereby demonstrating a much stronger product band than the wild-type holoenzyme protein, and the polymerization product could be efficiently inhibited by heparin (fig. S1).

Overall architecture of F8-A22-E4 polymerase holoenzyme

To capture the replicating conformation of the MPXV F8-A22-E4 polymerase holoenzyme, we incubated the 3′-H modified primer-template DNA and the exonuclease-deficient polymerase holoenzyme in the reaction buffer with deoxythymidine triphosphate (dTTP) substrate. We then prepared cryo–electron microscopy (cryo-EM) samples using a graphene grid to avoid preferential orientation observed with ordinary grids. The holoenzyme–DNA complex was resolved to ~2.8 Å (figs. S2 and S3). The EM map shows the key structural features of all proteins and DNA elements (fig. S4). Although the density of the 5′-end template was weak, we traced the main chains using the unsharpened EM density map to demonstrate the template entry channel (see below).

The structure of the holoenzyme–DNA complex contains one F8, one A22, one E4, and the primer-template DNA, as well as an incoming dTTP substrate (Fig. 1). F8, A22, and E4 form pairwise interactions with each other (Fig. 1).

The F8 structure can be traced for 1004 residues, except for the last two residues, and the classical N-terminal domain (NTD), exonuclease domain (Exo), palm domain, fingers domain, and thumb domain were observed in a closed conformation (Fig. 1). Five "poxvirus-specific" insertion regions in MPXV F8 can also

be observed, like those inserts seen in the VACV E9 structure (*16*) (Fig. 1 and fig. S5, A and B).

The 218-residue MPXV E4 resembles the VACV D4 structure (*25*) (Fig. 1A and fig. S5, C and D). The 426-residue MPXV A22 structure can be divided into three domains: the A22 NTD, middle domain (Mid), and C-terminal domain (CTD) (Fig. 1A and fig. S5E). When we aligned the Mid of A22 on the online Dali server, we found that it shows high structural similarities with the African swine fever virus (ASFV) DNA ligase and the bacteriophage T4 DNA ligase (*26*, *27*). The Mid can be further divided into two subdomains: an adenylation domain (OD), which mostly resembles the OD of ASFV DNA ligase (fig. S5F), and an OB-fold domain (OB), which mostly resembles the OB of *Thermus filiformis* DNA ligase (fig. S5G). Further enzymatic assay showed that the MPXV polymerase holoenzyme did not possess ordinary ligase activity similar to that of T4 ligase (fig. S6, A and B). Compared with the structures of adenylation domains from the T4 and ASFV ligases, the putative ligase active site of the A22 Mid is replaced by hydrophobic and negatively charged residues, which may prevent the binding of ATP (fig. S6C). As the Mid of A22 lacks the essential DNA binding domain, A22 may comprise a degenerative ligase domain acting simply as a flexible linker.

The MPXV DNA polymerase holoenzyme is stabilized by pairwise interactions between the F8, A22, and E4 subunits (fig. S7). A22 acts as a bridge to bind E4 and F8 via the A22 NTD and CTD, respectively. The interactions are almost identical to those of their VACV counterparts (fig. S8) (*21*–*24*). A previous study proposed a VACV polymerase holoenzyme model with an elongated shape of the A20-D4 cofactor, leading to a ~150-Å distance between the E9 polymerase active site and the D4 DNA binding site (*28*). However, in our replicating MPXV holoenzyme structure, the A22-E4 cofactor folds back, and E4 directly interacts with the Exo domain of F8 at two sites, one where Trp^{36} and Arg^{39} of E4 form hydrogen bonds and hydrophobic interactions with Phe^{179} and Leu^{278} from F8 Exo (fig. S7F), and another where Asn^{165} of E4 forms a hydrogen bond with Asn^{303} from F8 Exo (fig. S7G).

Primer-template DNA recognition by the polymerase complex

The structure of MPXV polymerase holoenzyme–DNA complex contains 22-nt DNA in the template strand, 14-nt DNA in the primer strand, and the incoming dTTP, as well as a magnesium ion that may serve as catalytic ion near the active site (Fig. 2, A and B). The double-stranded primer-template DNA binds in a groove formed between the palm and thumb domains of F8, and the single-stranded 5′ extension of the template strand probably passes through a channel formed by the NTD and

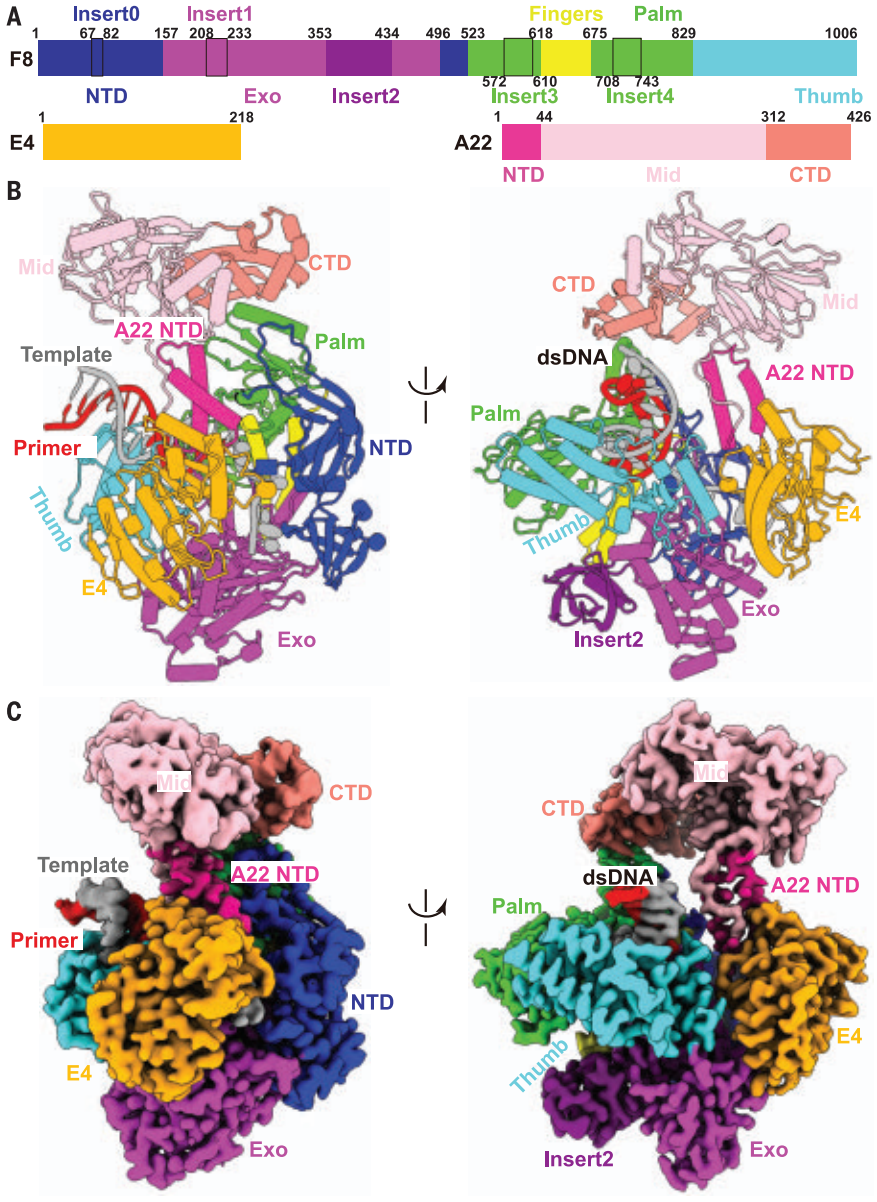


Fig. 1. Overall structure of the replicating MPXV DNA polymerase holoenzyme. (**A**) Schematic diagrams of the domain architecture of MPXV DNA polymerase F8 and processivity factors (A22 and E4). The F8 can be divided into five domains: NTD, blue; Exo, magenta; palm, green; fingers, yellow; thumb, cyan. Compared to other B-family polymerases, F8 contains five inserted elements in which the largest one was named as insert2 (purple), and the other four small inserts are indicated as rectangles. A22 is colored by domains: NTD, deep pink; Mid, pink; CTD, salmon. E4, orange; template strand, gray; primer strand, red. (**B** and **C**) Atomic model and cryo-EM density map of the replicating MPXV DNA polymerase holoenzyme. The structures were colored by domains, as depicted in (A).

Exo of F8 and the E4 subunit in an orientation perpendicular to the DNA duplex (Fig. 2, C and D, and fig. S9). The template DNA has 12 unpaired nucleotides at the 5′ end, but only two of them are well-ordered with a defined base structure. For the remaining 10 unpaired bases, we can only trace partial phosphate-ribose backbones (C6 to T11) because of the weak EM density (Fig. 2, C and D).

Upon primer-template DNA binding, F8 polymerase undergoes conformational changes, a common feature of the B-family DNA polymerases. Comparison of the MPXV F8 in this holoenzyme–DNA complex structure with the

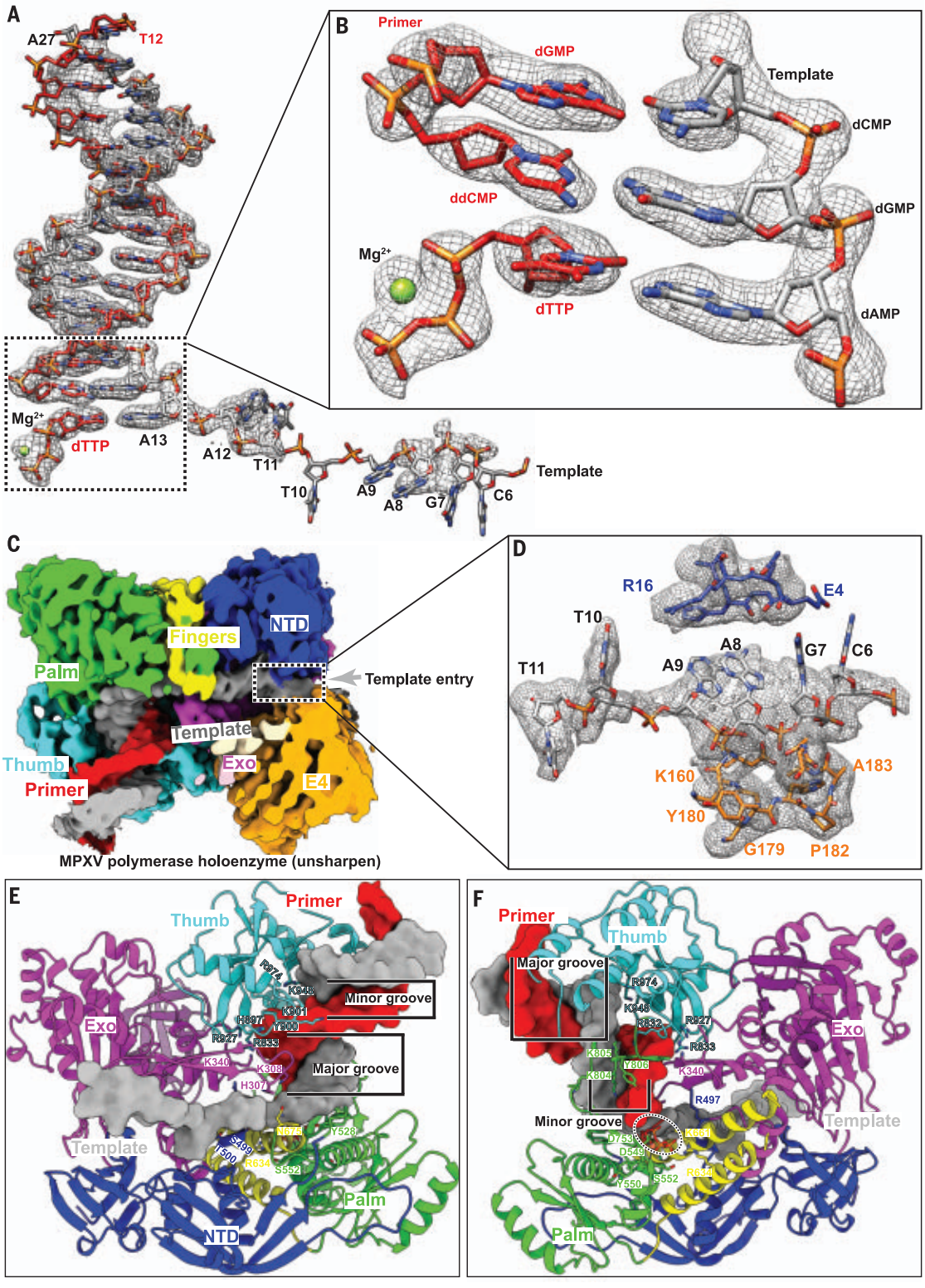


Fig. 2. The interactions between DNA and MPXV polymerase. (**A**) Sharpened EM densities and atomic models of dsDNA. (**B**) Enlarged view of the sharpened EM densities and atomic models of dsDNA in the active site of polymerase. (**C**) The cut-off view of unsharpened EM density map for MPXV polymerase holoenzyme in complex with the DNA, revealing the consecutive density of template DNA. (**D**) Enlarged view of the unsharpened densities and supposed atomic models of the template DNA in the template entry channel. The region (C6 to T11) was built using this unsharpened map. The remaining 5′-terminal five bases of template DNA were invisible, which reflects an inherent flexible conformation of the 5′-terminal unpaired region of the template strand. (**E** and **F**) The primary interfaces between F8 and DNA. The F8 mainly interacted with the minor groove of primer-template DNA, with only a few contacts to the major groove contributed by the residues of exonuclease domain. The primer-template DNA is shown in surface representation calculated from the atomic model, and F8 is shown in cartoon representation.

VACV apo E9 structure, which has high sequence identity, shows that the fingers domain rotates toward the palm domain by ~17° in the replicating state (fig. S9). This rotation drags the positively charged Arg^{634} and Lys^{661} of the fingers domain closer to the active site, where they can interact with the triphosphate group of incoming dNTP. The rotated fingers domain interacts with the Exo, and this interaction further stabilizes the closed conformation of the fingers domain. Moreover, the thumb domain also makes a distinct rotation to wrap around the primer-template DNA duplex on its minor groove side (fig. S9). The DNA duplex is accommodated in a positively charged groove of the thumb domain, as observed in other B-family DNA polymerases (fig. S10).

The modeled double-stranded DNA helix is formed by 14 base pairs from the primer-template DNA and maintains a B-form conformation (Fig. 2A and fig. S11). Extensive

protein-DNA interactions are observed between the primer-template DNA and F8, with a total of 47 residues from F8 directly participating in DNA binding (within 4.0 Å distance, 29 residues to the template strand, and 18 residues to the primer strand) (Fig. 2, fig. S11, and table S2). Most of the key residues involved in primer-template DNA binding are highly conserved among different B-family DNA polymerases. Protein-DNA interactions mainly involve the phosphodiester backbone of the DNA, with many interactions directly with the phosphate groups (fig. S11A). Interactions with the template-strand phosphates are largely hydrogen bonds to the main or side chains of 13 residues from the thumb, palm, fingers domains, Exo, and NTD of F8 (fig. S11, A to C), whereas the primer strand is bound by both electrostatic interactions and hydrogen bonds with the nine residues from the thumb, palm domains, and Exo of F8 (fig. S11, A, D, and E). There is little contact of F8 with the base pairs of primer-template DNA, except for one hydrogen bond interaction between R832 of the thumb domain and the base of T22 from the primer strand, which may be important for stabilizing the B-form conformation of the DNA duplex (fig. S11, A and D). This is consistent with the fact that the enzymatic activity of F8 does not rely on a specific sequence during the elongation step. In addition, residue N675 of the fingers domain forms a hydrogen bond with the base of unpaired A12 from the template strand, and this interaction may be responsible for the kinking of the single-stranded 5′ extension of the template strand near the active site.

Interactions between the polymerase and the incoming nucleotide

Next to the 3′ terminus of the primer strand is the incoming dTTP, which binds to the active site of the polymerase in a manner analogous to that observed in the structures of other DNA polymerase complexes (Fig. 3A) (*29*). The incoming dTTP is accommodated in a groove formed by residues from the palm and fingers domains (Fig. 3B). The two highly conserved aspartate residues, D549 (in motif A) and D753 (in motif C), together with the triphosphate tail of dTTP, coordinate one divalent metal ion (assumed as magnesium, which has been added to the reaction buffer) (Fig. 3C). The triphosphate tail also interacts with the main chains of Y550, S552, and L553 from motif A, and the side chains of two positively charged R634 and K661 from the fingers domain (Fig. 3C).

The ribose of dTTP stacks on top of the phenyl ring of Y554 from motif A, in a manner similar to that previously observed with Y416 in the ternary complex structure of RB69 polymerase (*30*) (Fig. 3C). There would be a steric clash between the 2'OH of ribonucleotides and Y554, hence providing a "steric gating" effect

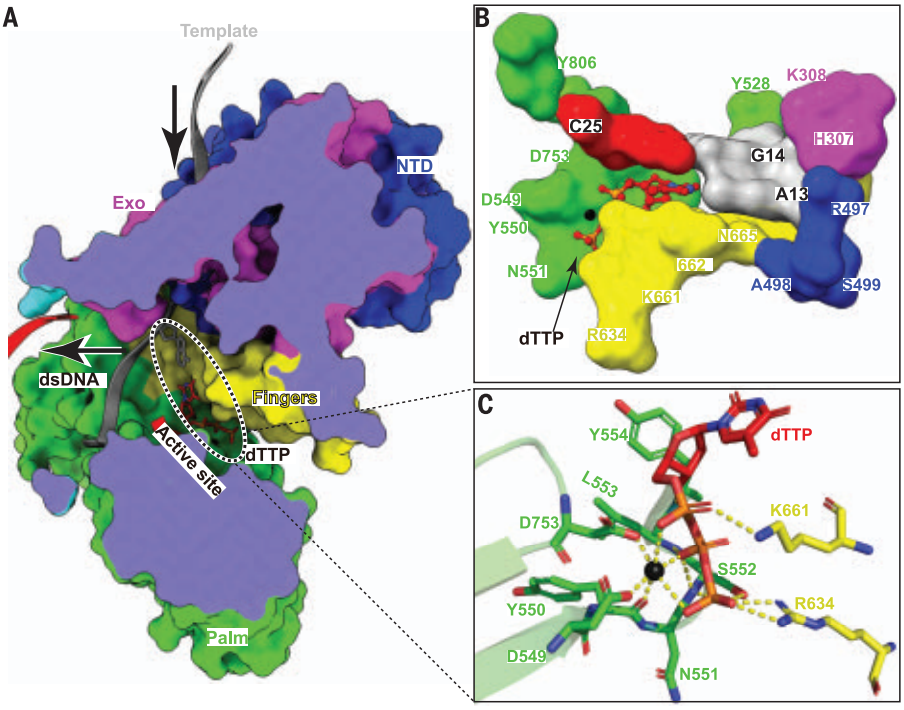


Fig. 3. Recognition of the incoming dTTP. (**A**) Cut-off view of the F8 protein, which is shown in surface representation to reveal the inner active site. The F8 is colored by domains as in Fig. 1; template strand, gray; primer strand, red. (**B**) The binding pocket of the incoming dTTP. It is formed by the fingers domain, palm domain, and upper base pair. The incoming dTTP is shown as a ball-and-stick model, and the residues of F8 and the upper base pair are shown in surface representation. (**C**) Interaction details between F8 and the incoming dTTP. The key residues are shown as sticks and colored as corresponding domains. The magnesium ion is depicted as a black sphere. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

to select dNTP as the substrate. This similar "steric gating" effect has also been described in other DNA polymerases and was first proposed in HIV-1 reverse transcriptase (*31*–*33*).

Operation mode of processivity cofactor

For B-family DNA polymerases, proliferating cell nuclear antigen (PCNA) or PCNA-like proteins are required for high processivity. However, for poxviruses, including MPXV and VACV, no homologous PCNA-like proteins have been identified in the viral genome. Instead, the poxvirus-specific A22-E4 heterodimer is responsible for the high processivity of DNA replication.

A primer-extension assay using a 60-nt template DNA in the presence of heparin, which can trap the dissociated DNA polymerase from the primer-template DNA to guarantee a single-turnover reaction, showed that the F8 polymerase alone dissociated from the primer-template DNA after incorporating less than 14 nt, whereas the F8-A22-E4 holoenzyme was able to generate full-length 60-nt products with few abortive ones (Fig. 4A). We then demonstrated that the addition of the A22-E4 heterodimer conferred processivity to F8 in a concentration-dependent manner (Fig. 4A). When the molar ratio of A22-E4 to F8 was 1:1, corresponding to the stoichiometry of the polymerase holoenzyme, the amount of the full-length product was almost the same as that of the product generated by the preassembled F8-A22-E4 polymerase holoenzyme (Fig. 4A). This indicates that the isolated A22-E4 and F8 can be efficiently assembled into functional holoenzymes to perform processive DNA synthesis. Moreover, we performed alanine scanning of critical residues responsible for the interaction between E4 and F8 and found that R39A and N165A substitutions of E4 showed minor effect, W36A reduced the synthesis of full-length products, whereas the W36A/R39A and W36A/R39A/N165A substitutions abolished the synthesis of full-length products (Fig. 4B). These results further confirmed the important function of the A22-E4 heterodimer in DNA replication processivity in a pure enzymatic reaction system.

As described above, the E4 cofactor interacts with the Exo of F8 polymerase, and together with the NTD of F8, they form a closed-ring channel to encircle the single-stranded template DNA (Fig. 4C). By contrast, in the yeast DNA polymerase complex (*29*), a representative of the other B-family DNA polymerases (fig. S12), the Exo and NTD form an open semicircular

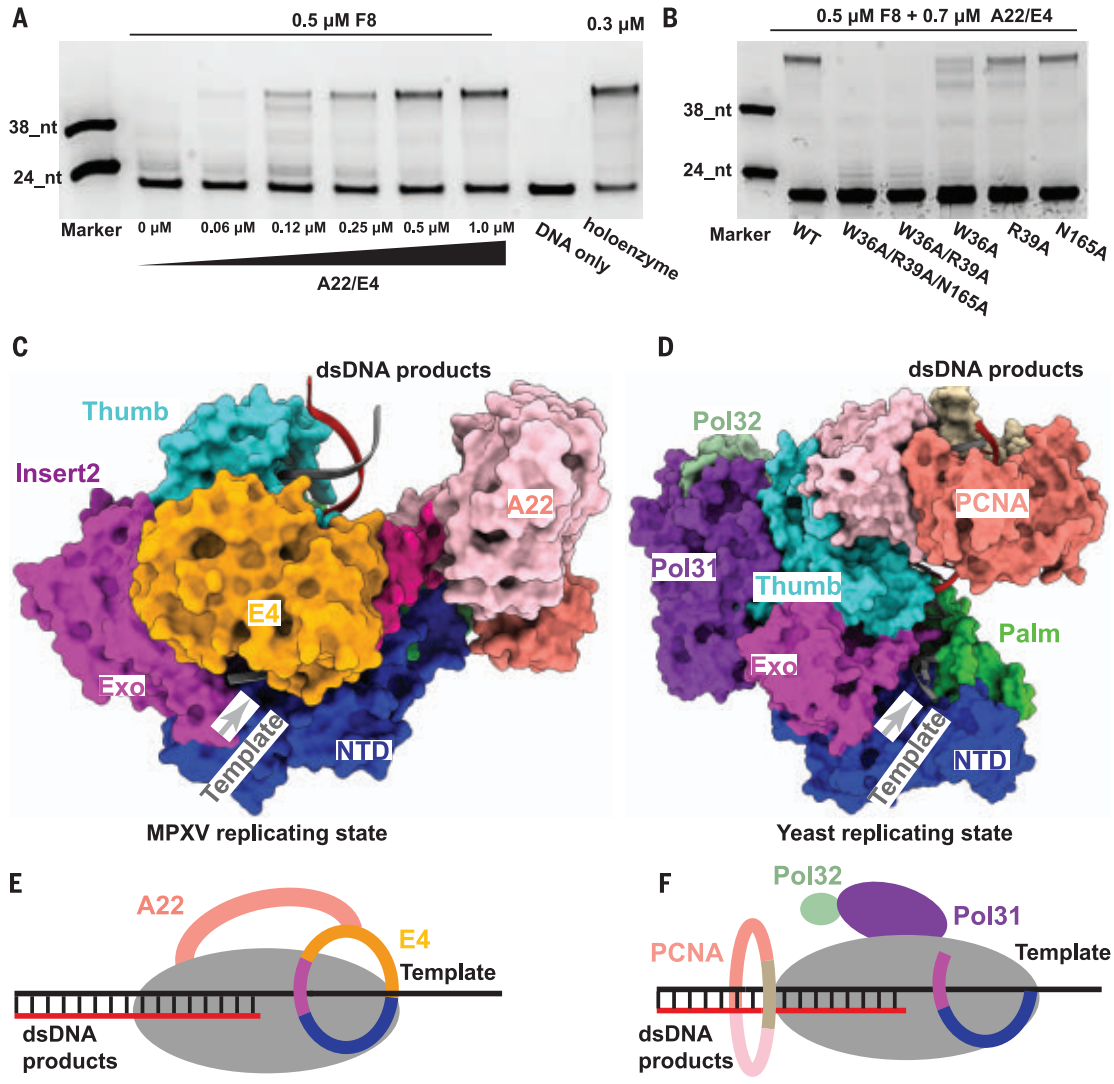


Fig. 4. The distinctive mechanism of A22-E4 heterodimer promoting processivity of MPXV DNA polymerase. (**A**) Enzymatic assay of A22-E4 heterodimer protein improving the DNA replication processivity of exonuclease-deficient F8 under single-turnover conditions. The F8 alone was shown to be distributive and failed to generate full-length products, whereas the addition of the A22-E4 complex promoted the yield of full-length DNA in a concentration-dependent manner. (**B**) Alanine scanning of critical residues of E4 responsible for the interaction with F8 was performed to examine the effects on the DNA replication processivity. The A22-E4 W36A mutant reduced the processivity activity, whereas the W36A/R39A and W36A/R39A/N165A mutants abolished the processivity activity. The F8 used in this assay was exonuclease deficient. (**C** and **D**) Comparison of the structures of the MPXV and yeast polymerases (PDB ID: 7KC0) in complex with their own processivity factors and DNA. Polymerases and processivity factors are shown in surface representation and colored by domains, as in Fig. 1, whereas the primer-template DNA strands are shown as a cartoon (template, gray; primer, red). The processivity of yeast polymerase was strengthened by trimeric PCNA to clamp the primer-template dsDNA. While in the structure of MPXV polymerase holoenzyme, A22-E4 heterodimer does not interact with dsDNA. Instead, E4 located on the template entry channel combined with NTD and Exo domains of F8 to form a forward clamp structure that would prevent the template strand disassociating from the polymerase complex during DNA replication. (**E** and **F**) Two binding modes of processivity factors with polymerases. The processivity factors bound with template in poxvirus function as a "forward sliding clamp" (E) or dsDNA products in eukaryotes as a "backward sliding clamp" (F).

channel to accommodate the single-stranded template DNA, and the trimeric PCNA ring encircles the template-product DNA duplex (Fig. 4D). This architectural difference between MPXV and yeast polymerase complexes is responsible for the different processivity mechanisms during DNA replication events. The MPXV DNA polymerase holoenzyme guarantees its high DNA replication processivity by encircling the single-stranded template DNA, and we propose that it functions as a "forward sliding clamp" (Fig. 4E); whereas the other B-family DNA polymerase complexes possess continuous DNA replication capacity by encircling the double-stranded template-product DNA helix that can be recognized as a "backward sliding clamp" (Fig. 4F).

Discussion

The interaction between E4 and F8 could generate a ring channel that encircles the single-stranded template DNA, which is proposed to be important for high DNA replication processivity. This processivity mechanism is different from that of other B-family DNA polymerases that utilize PCNA or PCNA-like proteins to encircle the product-template DNA duplex (*29*, *34*, *35*). The configuration of encircling the single-stranded template DNA probably allows the MPXV polymerase complex to perform continuous DNA replication by preventing template DNA disassociation from the polymerase holoenzyme. The efficiency

of polymerase holoenzyme assembly in the cellular environment requires further study, and moreover, it should be investigated whether the F8-A22-E4 polymerase holoenzyme has other active conformations (*28*).

Previous studies have implicated the distinctive role of the polymerase holoenzyme in DNA recombination, which involves both 3′-5′ exonuclease and DNA-joining activities (*36*–*38*). Although we identified DNA ligase-like domains for the Mid of the poxvirus-specific cofactor A22, the F8-A22-E4 polymerase holoenzyme did not have canonical ligase activity in our assay. However, we cannot rule out the possibility that A22 possesses ligase activity in a different state from the current holoenzyme structure or in an ATP-independent manner. Orthopoxviruses encode a DNA ligase that is not essential for virus replication but affects virulence and sensitivity to DNA-damaging agents (*39*). Further functional dissection of the middle part of A22 and its cooperation with viral and host ligases is needed.

Poxvirus replication processes use different replication models, including self-priming, primer-dependent, and recombination models (*15*). Poxviruses, including MPXV, have linear double-stranded DNA genomes, and the termini of the two DNA strands are connected to form a continuous polynucleotide chain (*40*). A rolling cycle mechanism has been proposed to replicate DNA in the form of unbranched head-to-tail concatemers, which would be resolved by a Holliday junction resolvase to produce unit genomes (*41*). The proposed "forward sliding clamp" mode can help interpret the self-priming model, as the processivity cofactor would facilitate robust continuous replication along the single-stranded template DNA unwound by the primase-helicase. Because there is also a possible presence of Okazaki fragments for *Orthopoxvirus* DNA replication (*42*), other replication models should also be studied. Moreover, the working mechanism of an intact replisome, including the F8-A22-E4 holoenzyme and other replication proteins, will be a fascinating area for future studies.

REFERENCES AND NOTES

1. H. Zhao *et al*., *China CDC Wkly.* **4**, 853–854 (2022).
2. J. G. Rizk, G. Lippi, B. M. Henry, D. N. Forthal, Y. Rizk, *Drugs* **82**, 957–963 (2022).
3. J. Kaler, A. Hussain, G. Flores, S. Kheiri, D. Desrosiers, *Cureus* **14**, e26531 (2022).
4. D. Focosi, F. Novazzi, A. Baj, F. Maggi, *Rev. Med. Virol.* **32**, e2392 (2022).
5. D. B. Di Giulio, P. B. Eckburg, *Lancet Infect. Dis.* **4**, 15–25 (2004).
6. E. M. Bunge *et al*., *PLOS Negl. Trop. Dis.* **16**, e0010141 (2022).
7. W. Tan, G. F. Gao, *China CDC Wkly.* **4**, 847–848 (2022).
8. N. Erez *et al*., *Emerg. Infect. Dis.* **25**, 980–983 (2019).
9. A. Vaughan *et al*., *Euro Surveill.* **23**, 1800509 (2018).
10. N. Chen *et al*., *Virology* **340**, 46–63 (2005).
11. A. M. McCollum, I. K. Damon, *Clin. Infect. Dis.* **58**, 260–267 (2014).
12. J. Isidro *et al*., *Nat. Med.* **28**, 1569–1572 (2022).
13. B. Moss, *Semin. Cell Dev. Biol.* **60**, 89–96 (2016).
14. M. D. Greseth, P. Traktman, *Annu. Rev. Virol.* **9**, 239–259 (2022).
15. B. Moss, *Cold Spring Harb. Perspect. Biol.* **5**, a010199 (2013).
16. N. Tarbouriech *et al*., *Nat. Commun.* **8**, 1455 (2017).
17. E. S. Stanitsa, L. Arps, P. Traktman, *J. Biol. Chem.* **281**, 3439–3451 (2006).
18. A. M. Druck Shudofsky, J. E. Silverman, D. Chattopadhyay, R. P. Ricciardi, *J. Virol.* **84**, 12325–12335 (2010).
19. M. W. Czarnecki, P. Traktman, *Virus Res.* **234**, 193–206 (2017).
20. K. A. Boyle, E. S. Stanitsa, M. D. Greseth, J. K. Lindgren, P. Traktman, *J. Biol. Chem.* **286**, 24702–24713 (2011).
21. C. Contesto-Richefeu *et al*., *Acta Crystallogr. F Struct. Biol. Commun.* **72**, 687–691 (2016).
22. W. P. Burmeister *et al*., *J. Biol. Chem.* **290**, 17923–17934 (2015).
23. C. Contesto-Richefeu *et al*., *PLOS Pathog.* **10**, e1003978 (2014).
24. B. Bersch, N. Tarbouriech, W. P. Burmeister, F. Iseni, *J. Mol. Biol.* **433**, 167009 (2021).
25. N. Schormann *et al*., *BMC Struct. Biol.* **7**, 45 (2007).
26. K. Shi *et al*., *Nucleic Acids Res.* **46**, 10474–10488 (2018).
27. Y. Chen *et al*., *Nat. Commun.* **10**, 387 (2019).
28. C. Sèle *et al*., *J. Virol.* **87**, 1679–1689 (2013).
29. F. Zheng, R. E. Georgescu, H. Li, M. E. O'Donnell, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30344–30353 (2020).
30. M. C. Franklin, J. Wang, T. A. Steitz, *Cell* **105**, 657–667 (2001).
31. A. P. Silverman, S. J. Garforth, V. R. Prasad, E. T. Kool, *Biochemistry* **47**, 4800–4807 (2008).
32. G. J. Klarmann *et al*., *Biochemistry* **46**, 2118–2126 (2007).
33. G. Gao, M. Orlova, M. M. Georgiadis, W. A. Hendrickson, S. P. Goff, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 407–411 (1997).
34. Y. Shamoo, T. A. Steitz, *Cell* **99**, 155–166 (1999).
35. C. Lancey *et al*., *Nat. Commun.* **11**, 1109 (2020).
36. D. B. Gammon, D. H. Evans, *J. Virol.* **83**, 4236–4250 (2009).
37. M. D. Hamilton, A. A. Nuara, D. B. Gammon, R. M. Buller, D. H. Evans, *Nucleic Acids Res.* **35**, 143–151 (2007).
38. M. D. Hamilton, D. H. Evans, *Nucleic Acids Res.* **33**, 2259–2268 (2005).
39. S. M. Kerr *et al*., *EMBO J.* **10**, 4343–4350 (1991).
40. O. Faye *et al*., *Lancet Infect. Dis.* **18**, 246 (2018).
41. N. Sklenovská, in *Animal-Origin Viral Zoonoses*, Y. S. Malik, R. K. Singh, K. Dhama, Eds. (Springer Singapore, Singapore, 2020), pp. 39-68.
42. M. Esteban, J. A. Holowczak, *Virology* **78**, 57–75 (1977).

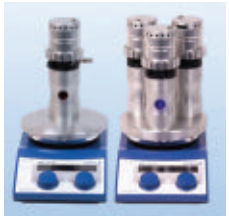
ACKNOWLEDGMENTS

We thank X.L., Y.M., and X.B. at the cryo-EM Center, Shanxi Academy of Advanced Research and Innovation, for their technical support of the cryo-EM data collection. **Funding:** National Key Research and Development Program of China grant 2021YFC2300700 (Y.S.); Strategic Priority Research Program of CAS grant XDB29010000 (Y.S., G.F.G.); National Natural Science Foundation of China (NSFC) grants 82241076, 81871658, 32192452, and 32100119 (Y.S., Q.P.); The Youth Innovation Promotion Association of CAS grant Y201921 (Y.S.).
Author contributions: Conceptualization: Y.S., G.F.G., Q.P. Investigation: L.K., Q.P., Y.X., J.Q., H.W. Visualization: Q.P., Y.X., L.K. Funding acquisition: Y.S., G.F.G., Q.P. Supervision: Y.S., G.F.G. Writing – original draft: Y.S., Q.P., G.F.G. Writing – review and editing: Y.S., Q.P., G.F.G. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Cryo-EM map is available in the Electron Microscopy Data Bank with code EMD-34731. Structural model is available in the Protein Data Bank (PDB) with accession code 8HG1. All materials are available from the authors on reasonable request with materials transfer agreements (MTAs).
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. In the interest of rapid dissemination of results with immediate public health relevance, the author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade6360
Materials and Methods
Figs. S1 to S12
Tables S1 and S2
References (*43*–*53*)
MDAR Reproducibility Checklist

Submitted 29 August 2022; accepted 5 December 2022
Published online 15 December 2022
10.1126/science.ade6360



Photochemical Reactor System
Asynt announces the LightSyn Lighthouse–a new photochemical reactor system in their LightSyn range, developed in response to customer demand for higher reaction yields, ease of use, high operational safety, and competitive pricing. The LightSyn Lighthouse uses new technology to channel light through a quartz rod directly into the reaction medium, maximizing power intensity while keeping photon flux even throughout. Using this system, the number of photons lost over distance is significantly reduced in comparison to that of a more typical photoreactor, therefore maximizing the light's interaction with the sample. The result is faster chemistry and higher reaction yields. Designed to be a useful tool for any lab looking to branch out into larger scale benchtop photochemistry, the LightSyn Lighthouse is simple to set up and to use, with built-in safety features that eliminate risk of UV light exposure. Easily adapted to suit different sized tubes/vials and round bottom flasks, the LightSyn Lighthouse is an effective and versatile system for scaling your experiments.
Asynt
For info: +44-0-1638-781-709
www.asynt.com/product/lightsyn-lighthouse

3'3'-cGAMP ELISA Kit
Cayman Chemical's 3'3'-cGAMP ELISA Kit is a competitive ELISA that can be used to quantify levels of the bacterial cyclic dinucleotide second messenger 3'3'-cyclic GMP-AMP (3'3'-cGAMP) in bacterial and mammalian cell lysates and cell supernatants. The assay uses a highly specific monoclonal antibody for excellent sensitivity (LLOD = 26 pM) and selectivity, has an assay range of 78-10,000 pM, and has been validated by LC-MS/MS to ensure accurate, reliable results. This kit was developed in collaboration with the nucleotide experts at Biolog Life Science Institute, and allows you to assay 24 samples in triplicate or 36 samples in duplicate in under 3 h with a plate-based colorimetric readout.
Cayman Chemical
For info: +1-800-364-9897
www.caymanchem.com/33cGAMP

Reagent Reservoirs
Manufactured to the ANSI/SLAS standard, the range of reagent reservoirs from Porvair Sciences are designed for simple integration into any automated liquid handling system. Moulded from high-purity polypropylene in a Class 10,0000 cleanroom environment, Porvair reagent reservoirs offer chemical compatibility with most organic solvents, acids, and bases. High heat resistance allows the reagent reservoirs to be autoclaved clean. With a choice of 20 working configurations, as well as liquid volumes, an optimized reagent reservoir is available to suit liquid handling applications using 8- or 12-channel pipettes right through to 96- and 384-tip automated pipetting systems.
Porvair Sciences
For info: +1-856-696-3605
www.microplates.com/microplates/products/plate-type/reservoir

Human Cytokine Kit
Gyros Protein Technologies AB introduces Gyrolab Human Cytokine Kit Reagents, the first in a range of biomarker kits. These include five single analyte biomarker kit reagents for the quantification of human inflammatory cytokines: human TNF-a, IFN-gamma, IL-6, IL-10, and IL-4. The new Gyrolab Human Cytokine Kit Reagents are optimized for use on all Gyrolab systems, offer improved assay performance, and meet key customer needs to face the biomarker assay development challenges encountered during drug development programs and regulatory studies. These kits enable high-quality data to be generated through well-characterized kit reagents and standards, while providing a high sensitivity and wide dynamic range to cover the full spectrum of cytokine levels seen in disease states or pharmacodynamic studies. The Gyrolab Biomarker Kits offer convenience and an expedited analytical solution for biomarker quantification by removing the need for assay development. High throughput and automated analyses reduce manual intervention, enable short turnaround times, and accelerate data-driven decisions.
Gyros Protein Technologies
For info: +46-0-18-56-63-00
www.gyrosproteintechnologies.com

GMP-grade Human AB Serum for Cell and Gene Therapies
BioIVT supplies Good Manufacturing Practice (GMP) grade human AB serum, a cell culture supplement to enhance the development and manufacturing of cell and gene therapies. Human AB serum supports in vitro cell expansion by providing many of the growth factors, vitamins, nutrients, trace elements, and transport factors found in an in vivo environment. By helping to produce biologically relevant conditions, it allows many primary human cells to grow at a faster rate. This is particularly important for autologous therapies that require using a person's own cells or tissues, which may divide at a slower rate due to chemotherapy. BioIVT's GMP-grade human AB serum is produced using the same optimized protocols and enhanced regulatory oversight of source material collection, production, and processing as its research-use only material. But it also has additional viral testing—hepatitis A (plasma-derived only), hepatitis B core antibody, hepatitis E, and parvovirus B19 (plasma-derived only)—and associated documentation. BioIVT is offering both plasma-derived and off-the-clot serum collected from healthy, male AB blood type donors. All products are sterile, filtered, and ready for cell culture.
BioIVT
For info: 516-483-1196
https://bioivt.com

COVID-19 Omicron Research Products
AMS Biotechnology (AMSBIO) introduces further tools for SARS-CoV-2 researchers to help decipher the COVID-19 B.1.1.529 variant. AMSBIO's easy-to-perform Omicron assay kits come in a variety of read-out detection methods including TR-FRET, fluorogenic, colorimetric, and chemiluminescence. These new kits are ideal for use in validation studies and suitable for rapid high-throughput screening of SARS-CoV-2 inhibitors. Expressing B1.1.529 S protein, Omicron pseudoviruses from AMSBIO are invaluable tools for measuring inhibitory molecules when used in tandem with ACE2 recombinant cell lines or for measuring the activity neutralizing antibodies. AMSBIO's novel Omicron lentiviruses contain either luciferase or eGFP reporter genes that provide convenient readout of transduction. Also new to the market are a range of Omicron variant recombinant proteins, including nucleocapsid protein and 3CL protease and spike trimer (S1+S2), developed for the study of enzyme kinetics, screening inhibitors, and selectivity profiling.
AMS Biotechnology
For info: +1-617-945-5033
www.amsbio.com/sars-cov-2-spike-mutants

Electronically submit your new product description or product literature information! Go to www.science.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

UNIVERSITÉ DE GENÈVE

The Faculty of Medicine of the University of Geneva is seeking to fill a position of:

TENURE TRACK ASSISTANT PROFESSOR IN PSYCHIATRIC NEUROSCIENCE

CHARGE: This is a full-time position in psychiatric neuroscience to conduct translational and/or fundamental research relevant to clinical psychiatry. Candidates performing research using neuroimaging/neurophysiology approaches and/or molecular/cellular techniques using rodent models and/or organoid models to tackle translational questions in psychiatric neuroscience are of particular interest.

The candidate is expected to be involved in teaching neurosciences and psychiatry where appropriate, as well as in the supervision of academic master's and doctoral theses. He/she is also expected to take on management and organisational tasks at the Department of Psychiatry and at the Faculty level.

REQUIREMENTS:
PhD, MD-PhD or equivalent degree.
Experience in direction of research and skills for teaching.
Good knowledge of French and clinical experience in psychiatry would be an asset.
Publications in leading peer-reviewed journals.

ENVIRONMENT:
The successful candidate will be member of the Department of Psychiatry and affiliated member of the Department of Basic Neurosciences. Furthermore, he/she will be member of the Synapsy Center for Mental Health (**https://www.unige.ch/medecine/synapsycentre/en/**).

STARTING DATE: December 1st 2023 or according to agreement.

Mandatory online registration before **February 28th 2023** at: **http://www.unige.ch/academ**
Additional information may be obtained from: **viviane.burghardt@unige.ch**

The University is an equal opportunities employer and particularly welcomes applications from women

Science Careers helps you advance your career. Learn how!

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.

Visit ScienceCareers.org today — all resources are free

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

SCIENCECAREERS.ORG

Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences

Established in partnership between the Chinese Academy of Sciences and the Shenzhen Municipal Government, the Shenzhen Institute of Advanced Technology (SIAT) is a newly-created university with an objective to become the world's preeminent institute for emerging science and engineering programs. SIAT is equipped with state-of-art teaching and research facilities and is dedicated to cultivating international, visionary, and interdisciplinary talents while delivering research support to pursue innovation-driven development.

SIAT is located in Shenzhen, also known as the "Silicon Valley of China," a modern, clean, and green city, well-known for its stunning architecture, vibrant economy, and its status as a leading global technology hub. SIAT is seeking applications for faculty positions of all ranks in the following academic programs: Computer Science and Engineering, Bioinformatics, Robotics, Life Sciences, Material Science and Engineering, Biomedical Engineering, Pharmaceutical Sciences, Synthetic Biology, Neurosciences, etc. SIAT seeks individuals with a strong record of scholarship who possess the ability to develop and lead high-quality teaching and research programs. SIAT offers a comprehensive benefits package and is committed to faculty success throughout the academic career trajectory, providing support for ambitious and world-class research projects and innovative, interactive teaching methods.

Further information:
https://www.siat-sz.edu.cn/

ScienceCareers online @sciencecareers.org

What's Your Next Career Move?

From networking to mentoring to evaluating your skills, find answers to your career questions on *Science* Careers

To view the complete collection, visit
ScienceCareers.org/booklets

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

UNIVERSITÉ DE GENÈVE

The Faculty of Medicine of the University of Geneva is seeking to fill a position of:

TENURE TRACK ASSISTANT PROFESSOR IN PSYCHIATRY

CHARGE: This is a full-time position of research in psychiatric neurosciences. Candidates performing research using neuroimaging or neurophysiology approaches to tackle translational questions with clinical populations are of particular interest.
The incumbent will undertake research in this area at the national and international levels and secure external funding.
The position will involve undergraduate and graduate teaching in Psychiatry and supervision of Masters' and Doctoral theses.
He/She will also take on administrative and organizational duties within the Department of Psychiatry and the Faculty of Medicine.

REQUIREMENTS:
Doctorate of Medicine (MD) or equivalent degree
Full postgraduate training in psychiatry is required
Previous teaching and independent research experience
Publications in leading international journals
Good knowledge of French

STARTING DATE: December 1st 2023 or according to agreement.
Mandatory online registration before February 28th 2023 at: **http://www.unige.ch/academ**
Additional information may be obtained from: **viviane.burghardt@unige.ch**

Women are encouraged to apply

YOUR NEXT BIG SCIENTIFIC DISCOVERY: A NEW JOB.
Find your next job at ScienceCareers.org
Whether you're looking to get ahead, get into, or just plain get advice about careers in science, there's no better or more trusted authority. Get the scoop, stay in the loop with Science Careers.
ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ScienceCareers online @sciencecareers.org

By **Keisha Hardeman**

Tired of persevering

"Have you always wanted to be a scientist?" I have been asked this question many times before. This time, it's from a photographer who is taking my picture to advertise a new diversity initiative at the institution where I'm a postdoc. He seems nice and sincere, and I answer "yes" and tell him about cell division of adherent cancer cells. Still, his question stirs an ever-present uneasiness. Are other people asked this so frequently? I suspect I wouldn't hear the question as often if I were white or Asian. But I am Black. And I am forever being reminded, in ways big and small, that many think I don't belong in science or academia.

When I was a child growing up in a lower income African American family in the South, my parents told my brother and me we could be anything we wanted. On TV, I watched *The Cosby Show* featuring a Black family with a successful doctor dad and lawyer mom and saw young Black people in college on *A Different World*. I was completely in love with the natural world and won awards at grade school science fairs. I participated in state and national math and science competitions. In high school I took every advanced science class my school offered and was recruited into a research laboratory as part of early efforts to bring minorities into science. Surely science was where I belonged.

"The scientific challenges ... are nowhere near as discouraging as the systemic racism."

It was only when I started college at a predominantly white institution that I began to question whether I was a fool for thinking someone like me could be a scientist. I was usually the only Black student in my classes, and I had no Black professors. Fellow students excluded me from study groups. A professor even tried to kick me out of an exam, assuming I couldn't possibly be in his class. Thankfully, my family had faith in me and helped me remember that I am smart and capable. Even when I developed a mental health disability, their support made my aspirations seem viable.

After completing my bachelor's degree I went on to earn two professional medical certifications, and I finally found some much-needed encouragement from within the academic world when several wonderful white women professors readily accepted and nurtured me. Their mentorship made me want to become a professor myself—which meant pursuing a Ph.D.

Unfortunately, graduate school brought a return to the challenges and exclusion of my undergrad experience. I constantly heard about everything I was doing wrong and nothing about what I did well. If I had not already had therapy for my chronic depression and anxiety, I might have quit—or worse, followed the same path as two bright colleagues who died of suicide as a result of working in toxic environments, abused and mistreated.

The scientific challenges I encounter are nowhere near as discouraging as the systemic racism I've encountered in academia. I've had supposed colleagues critique my efforts within the first 30 seconds of meeting me, talk around and over me, push my name down the author list for no justifiable reason, and tell me I only earned awards because I'm Black. I've been misnamed Henderson, Hightower, Hardaway, and other "Black-sounding" names, including once in a departmental newsletter, prompting an uncomfortable email to the chair: "The newsletter looks great ... I wanted to point out my name is misspelled. ... Actually, that's just not my name." I've experienced treatment so harsh I've filed a complaint with the U.S. Equal Employment Opportunity Commission.

Despite the barriers I've encountered on my path—and in part because of them—I still dream of becoming a professor. But I'm tired of having to build resiliency, work on self-care strategies, and mentally process microaggressions that don't feel "micro" at all. I—and other Black scientists—want to bloom where we are planted, just like everyone else.

In my training and career across seven institutions in three states and abroad, I have only been taught by three Black professors. It would have made such a difference to have seen more faculty who looked like me. I hope I can make that difference for future students. ■

Keisha Hardeman is a postdoc at the University of Texas Southwestern Medical Center. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

YOUR RESEARCH HAS ITS REWARDS.

Submit an essay by 3/15/23 to enter to win $25K.

The prestigious international *Science* & PINS Prize is awarded annually to outstanding early-career researchers in the emerging field of neuromodulation. Findings already show improved outcomes in several neurological disorders. To enter, just write a 1000-word essay about your research performed in the last three years. The Grand Prize winner will be presented a US$25,000 award, as well as have their essay published in *Science* magazine and on *Science* Online. A runner-up will have their essay published online. For additional inquiries, email SciencePINSPrize@aaas.org.

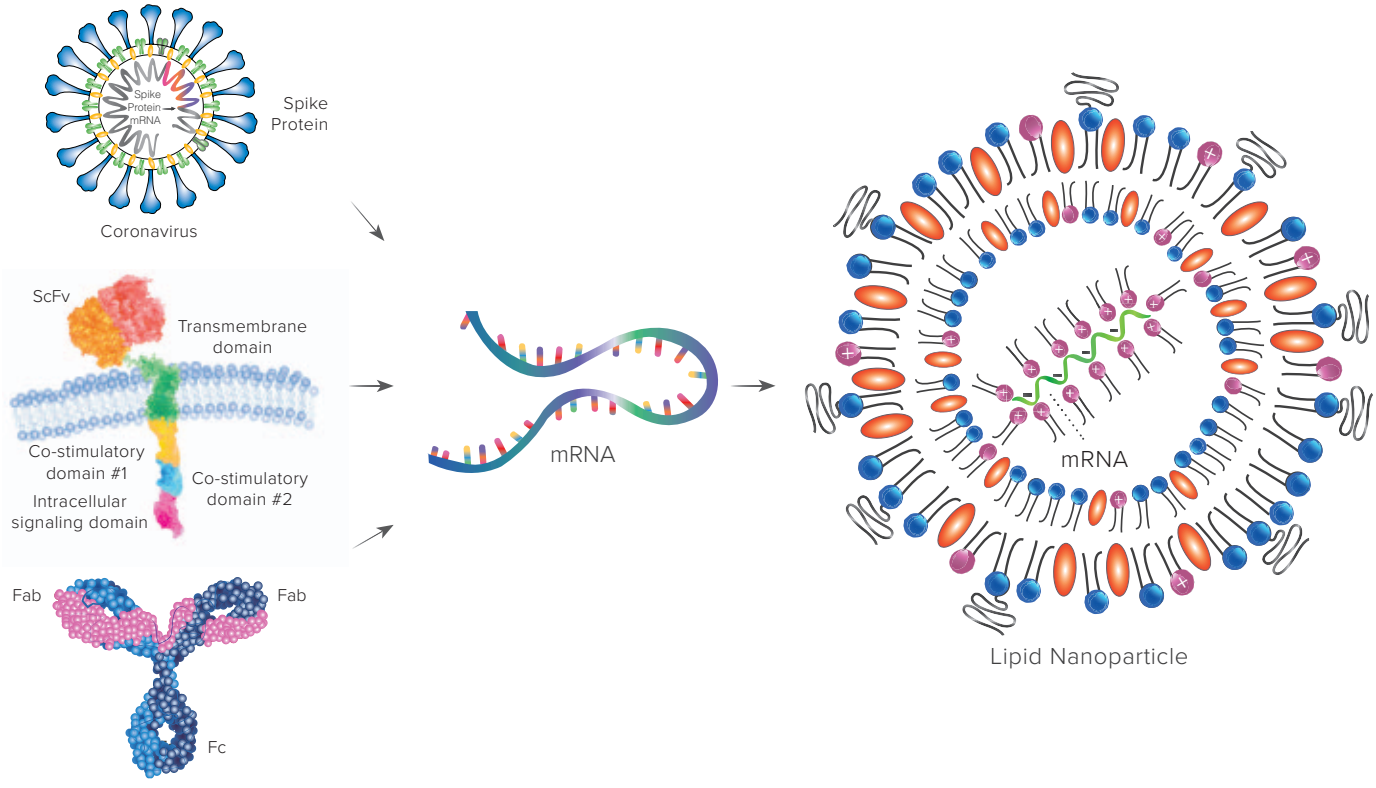
SCIENCE-PINS PRIZE FOR NEUROMODULATION

SCIENCE-PINS

SCIENCE.ORG/PINS

mRNA-LNP Products & Engineering

Providing LNP Solutions for Therapeutic Research



ProMab has expanded our product and service platform to offer ready-to-use mRNA lipid nanoparticles to accelerate your research. Lipid nanoparticles offer many benefits compared to other gene delivery methods and has far reaching applications in a variety of fields in therapeutic research. With strict quality control, unique encapsulation and formulations, and our expertise in antibody and cell therapy, we can deliver and validate many different genes of interest to help your research.

Current mRNA-LNP Products

Potential Applications	Gene of Interest	
• Viral Vaccine Development	• GFP	• MAGEA3/MHC protein
• Cancer Immunotherapies	• COVID-19 spike protein (Alpha variant)	• Anti-CD19 chimeric antigen receptor
• Immune Cell Engineering	• COVID-19 spike protein (Omicron variant)	• Anti-BCMA chimeric antigen receptor
• Protein Replacement Therapies	• COVID-19 spike protein (Delta variant)	• Anti-HER2 chimeric antigen receptor
• Cell and Gene Immunotherapies	• COVID-19 nucleocapsid protein	• EpCam/CD3 bispecific antibody
• Antibody Therapeutics	• COVID-19 RBD protein	

All products are for research use only

Discover more | www.promab.com

ProMab Biotechnologies, Inc.
2600 Hilltop Dr, Building B, Richmond, CA 94806
1.866.339.0871 | info@promab.com

SCAN ME